

STIC-Biotech/ChemLib

83917

From: Lacourciere, Karen  
Sent: Thursday, January 09, 2003 12:38 PM  
To: STIC-Biotech/ChemLib  
Subject: sequence search request 09/757,100

Please search SEQ ID NO:18 of 09/757,100 in both the commercial databases and the pending files (Interference)  
Thank-you!

*Karen A. Lacourciere Ph.D.*

CM1 11D09 GAU 1635

(703) 308-7523  
mailbox 11E12

Point of Contact:  
Beverly Shears  
Technical Info. Specialist  
CM1 1E05 Tel: 308-4994

STIC  
JAN 9 2003  
RECEIVED

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: \_\_\_\_\_  
Date Completed: \_\_\_\_\_  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:  
NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

**THIS PAGE BLANK (USPTO)**

## SEARCH REQUEST FORM

Requestor's Name: \_\_\_\_\_ Serial Number: \_\_\_\_\_  
Date: \_\_\_\_\_ Phone: \_\_\_\_\_ Art Unit: \_\_\_\_\_

### Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

### STAFF USE ONLY

Date completed: 01-13-03  
Searcher: Beverly e4774  
Terminal time: 23  
Elapsed time: \_\_\_\_\_  
CPU time: \_\_\_\_\_  
Total time: 23  
Number of Searches: \_\_\_\_\_  
Number of Databases: 22

Search Site  
\_\_\_\_ STIC  
\_\_\_\_ CM-1  
\_\_\_\_ Pre-S  
Type of Search  
\_\_\_\_ N.A. Sequence  
\_\_\_\_ A.A. Sequence  
\_\_\_\_ Structure  
\_\_\_\_ Bibliographic

Vendors  
\_\_\_\_ IG Suite  
\_\_\_\_ ☒ STN  
\_\_\_\_ Dialog  
\_\_\_\_ APS  
\_\_\_\_ Geninfo  
\_\_\_\_ SDC  
\_\_\_\_ DARC/Questel  
\_\_\_\_ ☒ Other CGN



09/757100

(FILE 'REGISTRY' ENTERED AT 10:55:53 ON 13 JAN 2003)  
L1 20 S TTTCAACCAGATGGTCATTC/SQSN

FILE 'HCAPLUS' ENTERED AT 10:57:04 ON 13 JAN 2003  
L2 7 S L1

L2 ANSWER 1 OF 7 HCAPLUS COPYRIGHT 2003 ACS

ACCESSION NUMBER: 2002:889441 HCAPLUS

DOCUMENT NUMBER: 137:381006

TITLE: Expressed sequence profiles and their use for  
the therapy and diagnosis of human lung cancer  
INVENTOR(S): Henderson, Robert A.; Wang, Tongtong; Watanabe,  
Yoshihiro; Johnson, Jeffrey C.; Retter, Marc W.;  
Durham, Margarita; Carter, Darrick; Fanger, Gary  
R.; Vedvick, Thomas S.; Bangur, Chaitanya S.;  
McNabb, Andria

PATENT ASSIGNEE(S): Corixa Corporation, USA

SOURCE: U.S. Pat. Appl. Publ., 82 pp., Cont.-in-part of  
U.S. Ser. No. 849,626.

CODEN: USXXCO

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 5

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 2002172952	A1	20021121	US 2001-902941	20010710
US 6504010	B1	20030107	US 2000-702705	20001030
US 2002168637	A1	20021114	US 2000-736457	20001213
US 2002197669	A1	20021226	US 2001-849626	20010503

PRIORITY APPLN. INFO.:

US 1999-419356	A2	19990630
US 1999-346492	A2	19991015
US 1999-466867	A2	19991217
US 1999-476300	A2	19991230
US 2000-519642	A2	20000306
US 2000-533077	A2	20000322
US 2000-546259	A2	20000410
US 2000-560406	A2	20000427
US 2000-589184	A2	20000605
US 2000-614124	A2	20000711
US 2000-651563	A2	20000829
US 2000-658824	A2	20000908
US 2000-671325	A2	20000926
US 2000-677419	A2	20001006
US 2000-702705	A2	20001030
US 2000-736457	A2	20001213
US 2001-849626	A2	20010503
WO 2000-US18061	A2	20000630

AB Nucleic acid and encoded protein compns. and methods for the therapy and diagnosis of cancer, particularly lung cancer, are disclosed. Thus, cDNAs encoding tumor-assocd. proteins were identified by conventional and PCR-based library subtraction, cDNA microarray anal., and real-time PCR using cDNA libraries isolated from lung adenocarcinoma, small cell lung carcinoma, squamous cell lung carcinoma, and lung neuroendocrine carcinoma vs. a pool of poly(A)+ RNA from 9 normal tissues. Illustrative compns. comprise one or more lung tumor polypeptides, immunogenic portions thereof,

**THIS PAGE BLANK (USPTO)**

09/757100

polynucleotides that encode such polypeptides, antigen presenting cell that expresses such polypeptides, and T cells that are specific for cells expressing such polypeptides. The disclosed compns. are useful, for example, in the diagnosis, prevention and/or treatment of diseases, particularly lung cancer.

IT 475864-34-7

RL: ANT (Analyte); BSU (Biological study, unclassified); DGN (Diagnostic use); PRP (Properties); THU (Therapeutic use); ANST (Analytical study); BIOL (Biological study); USES (Uses) (nucleotide sequence; expressed sequence profiles and their use for the therapy and diagnosis of human lung cancer)

L2 ANSWER 2 OF 7 HCAPLUS COPYRIGHT 2003 ACS

ACCESSION NUMBER: 2002:869469 HCAPLUS

DOCUMENT NUMBER: 137:364446

TITLE: Expressed sequence profiles and their use for the therapy and diagnosis of human lung cancer

INVENTOR(S): Wang, Tongtong; Bangur, Chaitanya S.; Lodes, Michael J.; Fanger, Gary R.; Vedvick, Thomas S.; Carter, Darrick; Retter, Marc W.; Mannion, Jane; Fan, Liqun; Wang, Aijun

PATENT ASSIGNEE(S): USA

SOURCE: U.S. Pat. Appl. Publ., 66 pp., Cont.-in-part of U. S. Ser. No. 702,705.

CODEN: USXXCO

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 5

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 2002168637	A1	20021114	US 2000-736457	20001213
WO 2001000828	A3	20010802	WO 2000-US18061	20000630
WO 2001000828	A2	20010104		
W:	AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM			
RW:	GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG			
EP 1196565	A2	20020417	EP 2000-945031	20000630
R:	AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, SI, LT, LV, FI, RO			
BR 2000012069	A	20020827	BR 2000-12069	20000630
US 6504010	B1	20030107	US 2000-702705	20001030
US 2002197669	A1	20021226	US 2001-849626	20010503
US 2002172952	A1	20021121	US 2001-902941	20010710
WO 2002004514	A2	20020117	WO 2001-US22058	20020710
W:	AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR,			

Searcher : Shears 308-4994

**THIS PAGE BLANK (USPTO)**



09/757100

TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ,  
MD, RU, TJ, TM  
RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH,  
CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE,  
TR, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD,  
TG

PRIORITY APPLN. INFO.:

US 1999-346492	A2 19990630
US 1999-419356	A2 19991015
US 1999-466867	A2 19991217
US 1999-476300	A2 19991230
US 2000-519642	A2 20000306
US 2000-533077	A2 20000322
US 2000-546259	A2 20000410
US 2000-560406	A2 20000427
US 2000-589184	A2 20000605
WO 2000-US18061	A2 20000630
US 2000-614124	A2 20000711
US 2000-651563	A2 20000829
US 2000-658824	A2 20000908
US 2000-671325	A2 20000926
US 2000-677419	A2 20001006
US 2000-702705	A2 20001030
US 2000-736457	A2 20001213
US 2001-849626	A2 20010503

AB Nucleic acid and encoded protein compns. and methods for the therapy and diagnosis of cancer, particularly lung cancer, are disclosed. Thus, cDNAs encoding tumor-assocd. proteins were identified by conventional and PCR-based library subtraction, cDNA microarray anal., and real-time PCR using cDNA libraries isolated from lung adenocarcinoma, small cell lung carcinoma, squamous cell lung carcinoma, and lung neuroendocrine carcinoma vs. a pool of poly(A)+ RNA from 9 normal tissues. Illustrative compns. comprise one or more lung tumor polypeptides, immunogenic portions thereof, polynucleotides that encode such polypeptides, antigen presenting cell that expresses such polypeptides, and T cells that are specific for cells expressing such polypeptides. The disclosed compns. are useful, for example, in the diagnosis, prevention and/or treatment of diseases, particularly lung cancer.

IT 475019-40-0

RL: ANT (Analyte); BSU (Biological study, unclassified); DGN (Diagnostic use); PRP (Properties); THU (Therapeutic use); ANST (Analytical study); BIOL (Biological study); USES (Uses)  
(nucleotide sequence; expressed sequence profiles and their use for the therapy and diagnosis of human lung cancer)

L2 ANSWER 3 OF 7 HCAPLUS COPYRIGHT 2003 ACS

ACCESSION NUMBER: 2002:332807 HCAPLUS

DOCUMENT NUMBER: 136:351185

TITLE: The contribution of 700,000 ORF sequence tags to the definition of the human transcriptome

AUTHOR(S): Camargo, Anamaria A.; Samaia, Helena P. B.;  
Dias-Neto, Emmanuel; Simao, Daniel F.; Migotto,  
Italo A.; Briones, Marcelo R. S.; Costa,  
Fernando F.; Nagai, Maria Aparecida;  
Verjovski-Almeida, Sergio; Zago, Marco A.;  
Andrade, Luis Eduardo C.; Carrer, Helaine;  
El-Dorry, Hamza F. A.; Espreafico, Enilza M.;  
Habr-Gama, Angelita; Giannella-Neto, Daniel;

**THIS PAGE BLANK (USPTO)**

Goldman, Gustavo H.; Gruber, Arthur; Hackel, Christine; Kimura, Edna T.; Maciel, Rui M. B.; Marie, Suely K. N.; Martins, Elizabeth A. L.; Nobrega, Marina P.; Paco-Larson, Maria Luisa; Pardini, Maria Ines M. C.; Pereira, Goncalo G.; Pesquero, Joao Bosco; Rodrigues, Vanderlei; Rogatto, Silvia R.; Da Silva, Ismael D. C. G.; Sogayar, Mari C.; Sonati, Maria De Fatima; Tajara, Eloiza H.; Valentini, Sandro R.; Alberto, Fernando L.; Amaral, Maria Elisabete J.; Aneas, Ivy; Arnaldi, Lilliane A. T.; De Assis, Angela M.; Bengtson, Mario Henrique; Bergamo, Nadia Aparecida; Bombonato, Vanessa; De Camargo, Maria E. R.; Canevari, Renata A.; Carraro, Dirce M.; Cerutti, Janete M.; Correa, Maria Lucia C.; Correa, Rosana F. R.; Costa, Maria Cristina R.; Curcio, Cyntia; Hokama, Paula O. M.; Ferreira, Ari J. S.; Furuzawa, Gilberto K.; Gushiken, Tsieko; Ho, Paulo L.; Kimura, Elza; Krieger, Jose E.; Leite, Luciana C. C.; Majumder, Paromita; Marins, Mozart; Marques, Everaldo R.; Melo, Analy S. A.; Melo, Monica; Mestriner, Carlos Alberto; Miracca, Elisabete C.; Miranda, Daniela C.; Nascimento, Ana Lucia T. O.; Nobrega, Francisco G.; Ojopi, Elida P. B.; Pandolfi, Jose Rodrigo C.; Pessoa, Luciana G.; Prevedel, Aline C.; Rahal, Paula; Rainho, Claudia A.; Reis, Eduardo M. R.; Ribeiro, Marcelo L.; Da Ros, Nancy; De Sa, Renata G.; Sales, Magaly M.; Sant'anna, Simone Cristina; Dos Santos, Mariana L.; Da Silva, Aline M.; Da Silva, Neusa P.; Silva, Wilson A., Jr.; Da Silveira, Rosana A.; Sousa, Josane F.; Stecconi, Daniella; Tsukumo, Fernando; Valente, Valeria; Soares, Fernando; Moreira, Eloisa S.; Nunes, Diana N.; Correa, Ricardo G.; Zalcborg, Heloisa; Carvalho, Alex F.; Reis, Luis F. L.; Brentani, Ricardo R.; Simpson, Andrew J. G.; De Souza, Sandro J.

CORPORATE SOURCE: Ludwig Institute for Cancer Research, Sao Paulo, 01509-010, Brazil

SOURCE: Proceedings of the National Academy of Sciences of the United States of America (2001), 98(21), 12103-12108

CODEN: PNASA6; ISSN: 0027-8424

PUBLISHER: National Academy of Sciences

DOCUMENT TYPE: Journal

LANGUAGE: English

AB Open reading frame expressed sequences tags (ORESTES) differ from conventional ESTs by providing sequence data from the central protein coding portion of transcripts. A total of 696,745 ORESTES sequences were generated from 24 human tissues and a subset of the data that correspond to a set of 15,095 full-length mRNAs used as a means of assessing the efficiency of the strategy and its potential contribution to the definition of the human transcriptome. It was estd. that ORESTES sampled over 80% of all highly and moderately expressed, and between 40% and 50% of rarely expressed, human genes. In the most thoroughly sequenced tissue, the breast, the 130,000

**THIS PAGE BLANK (USPTO)**

ORESTES generated are derived from transcripts from an estd. 70% of all genes expressed in that tissue, with an equally efficient representation of both highly and poorly expressed genes. In this respect, the capacity of the ORESTES strategy both for gene discovery and shotgun transcript sequence generation significantly exceeds that of conventional ESTs. The distribution of ORESTES is such that many human transcripts are now represented by a scaffold of partial sequences distributed along the length of each gene product. The exptl. joining of the scaffold components, by reverse transcription-PCR, represents a direct route to transcript finishing that may represent a useful alternative to full-length cDNA cloning. [This abstr. record is one of many records for this document necessitated by the large no. of index entries required to fully index the document and publication system constraints.].

IT 270220-29-6, GenBank AW992843 270220-33-2, GenBank AW992847

RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)

(nucleotide sequence; contribution of 700,000 ORF sequence tags to the definition of the human transcriptome)

L2 ANSWER 4 OF 7 HCAPLUS COPYRIGHT 2003 ACS

ACCESSION NUMBER: 2002:332795 HCAPLUS

DOCUMENT NUMBER: 136:351174

TITLE: The contribution of 700,000 ORF sequence tags to the definition of the human transcriptome

AUTHOR(S): Camargo, Anamaria A.; Samaia, Helena P. B.; Dias-Neto, Emmanuel; Simao, Daniel F.; Migotto, Italo A.; Briones, Marcelo R. S.; Costa, Fernando F.; Nagai, Maria Aparecida; Verjovski-Almeida, Sergio; Zago, Marco A.; Andrade, Luis Eduardo C.; Carrer, Helaine; El-Dorry, Hamza F. A.; Espreafico, Enilza M.; Habr-Gama, Angelita; Giannella-Neto, Daniel; Goldman, Gustavo H.; Gruber, Arthur; Hackel, Christine; Kimura, Edna T.; Maciel, Rui M. B.; Marie, Suely K. N.; Martins, Elizabeth A. L.; Nobrega, Marina P.; Paco-Larson, Maria Luisa; Pardini, Maria Ines M. C.; Pereira, Goncalo G.; Pesquero, Joao Bosco; Rodrigues, Vanderlei; Rogatto, Silvia R.; Da Silva, Ismael D. C. G.; Sogayar, Mari C.; Sonati, Maria De Fatima; Tajara, Eloiza H.; Valentini, Sandro R.; Alberto, Fernando L.; Amaral, Maria Elisabete J.; Aneas, Ivy; Arnaldi, Liliane A. T.; De Assis, Angela M.; Bengtson, Mario Henrique; Bergamo, Nadia Aparecida; Bombonato, Vanessa; De Camargo, Maria E. R.; Canevari, Renata A.; Carraro, Dirce M.; Cerutti, Janete M.; Correa, Maria Lucia C.; Correa, Rosana F. R.; Costa, Maria Cristina R.; Curcio, Cyntia; Hokama, Paula O. M.; Ferreira, Ari J. S.; Furuzawa, Gilberto K.; Gushiken, Tsieko; Ho, Paulo L.; Kimura, Elza; Krieger, Jose E.; Leite, Luciana C. C.; Majumder, Paromita; Marins, Mozart; Marques, Everaldo R.; Melo, Analy S. A.; Melo, Monica; Mestriner, Carlos Alberto; Miracca, Elisabete C.; Miranda, Daniela C.; Nascimento, Ana Lucia

**THIS PAGE BLANK (USPTO)**

09/757100

T. O.; Nobrega, Francisco G.; Ojopi, Elida P. B.; Pandolfi, Jose Rodrigo C.; Pessoa, Luciana G.; Prevedel, Aline C.; Rahal, Paula; Rainho, Claudia A.; Reis, Eduardo M. R.; Ribeiro, Marcelo L.; Da Ros, Nancy; De Sa, Renata G.; Sales, Magaly M.; Sant'anna, Simone Cristina; Dos Santos, Mariana L.; Da Silva, Aline M.; Da Silva, Neusa P.; Silva, Wilson A., Jr.; Da Silveira, Rosana A.; Sousa, Josane F.; Stecconi, Daniella; Tsukumo, Fernando; Valente, Valeria; Soares, Fernando; Moreira, Eloisa S.; Nunes, Diana N.; Correa, Ricardo G.; Zalberg, Heloisa; Carvalho, Alex F.; Reis, Luis F. L.; Brentani, Ricardo R.; Simpson, Andrew J. G.; De Souza, Sandro J.

CORPORATE SOURCE: Ludwig Institute for Cancer Research, Sao Paulo, 01509-010, Brazil

SOURCE: Proceedings of the National Academy of Sciences of the United States of America (2001), 98(21), 12103-12108

CODEN: PNASA6; ISSN: 0027-8424

PUBLISHER: National Academy of Sciences

DOCUMENT TYPE: Journal

LANGUAGE: English

AB Open reading frame expressed sequences tags (ORESTES) differ from conventional ESTs by providing sequence data from the central protein coding portion of transcripts. A total of 696,745 ORESTES sequences were generated from 24 human tissues and a subset of the data that correspond to a set of 15,095 full-length mRNAs used as a means of assessing the efficiency of the strategy and its potential contribution to the definition of the human transcriptome. It was estd. that ORESTES sampled over 80% of all highly and moderately expressed, and between 40% and 50% of rarely expressed, human genes. In the most thoroughly sequenced tissue, the breast, the 130,000 ORESTES generated are derived from transcripts from an estd. 70% of all genes expressed in that tissue, with an equally efficient representation of both highly and poorly expressed genes. In this respect, the capacity of the ORESTES strategy both for gene discovery and shotgun transcript sequence generation significantly exceeds that of conventional ESTs. The distribution of ORESTES is such that many human transcripts are now represented by a scaffold of partial sequences distributed along the length of each gene product. The exptl. joining of the scaffold components, by reverse transcription-PCR, represents a direct route to transcript finishing that may represent a useful alternative to full-length cDNA cloning. [This abstr. record is one of many records for this document necessitated by the large no. of index entries required to fully index the document and publication system constraints.].

IT 272727-43-2, GenBank BE092864

RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)

(nucleotide sequence; contribution of 700,000 ORF sequence tags to the definition of the human transcriptome)

L2 ANSWER 5 OF 7 HCAPLUS COPYRIGHT 2003 ACS

ACCESSION NUMBER: 2002:51525 HCAPLUS

DOCUMENT NUMBER: 136:129944

TITLE: Human nucleic acid and expressed protein

Searcher : Shears 308-4994

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)



09/757100

INVENTOR(S): compositions and methods for the therapy and diagnosis of lung cancer  
Wang, Tongtong; Watanabe, Yoshihiro; Henderson, Robert A.; Johnson, Jeffrey C.; Retter, Marc W.; Marnerakis, Margarita; Carter, Darrick; Fanger, Gary Richard; Vedvick, Thomas S.; Bangur, Chaitanya S.; McNabb, Andria; Wang, Aijun; Fanger, Neil; Switzer, Ann; McNeill, Patricia D.; Clapper, Jonathan D.  
PATENT ASSIGNEE(S): Corixa Corporation, USA  
SOURCE: PCT Int. Appl., 223 pp.  
CODEN: PIXXD2  
DOCUMENT TYPE: Patent  
LANGUAGE: English  
FAMILY ACC. NUM. COUNT: 5  
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2002004514	A2	20020117	WO 2001-US22058	20020710
W:	AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM			
RW:	GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG			
US 6504010	B1	20030107	US 2000-702705	20001030
US 2002168637	A1	20021114	US 2000-736457	20001213
US 2002197669	A1	20021226	US 2001-849626	20010503
PRIORITY APPLN. INFO.:			US 2000-614124	A 20000711
			US 2000-651563	A 20000829
			US 2000-658824	A 20000908
			US 2000-671325	A 20000926
			US 2000-677419	A 20001006
			US 2000-702705	A 20001030
			US 2000-736457	A 20001213
			US 2001-849626	A 20010503
			US 1999-419356	B2 19990630
			US 1999-346492	A2 19991015
			US 1999-466867	A2 19991217
			US 1999-476300	A2 19991230
			US 2000-519642	A2 20000306
			US 2000-533077	A2 20000322
			US 2000-546259	A2 20000410
			US 2000-560406	A2 20000427
			US 2000-589184	A2 20000605
			WO 2000-US18061	A2 20000630
AB	Compns. and methods for the therapy and diagnosis of cancer, particularly lung cancer, are disclosed. Illustrative compns. comprise one or more lung tumor polypeptides, immunogenic portions thereof, the cDNA polynucleotides that encode such polypeptides, antigen-presenting cell that expresses such polypeptides, and T cells that are specific for cells expressing such polypeptides.			

**THIS PAGE BLANK (USPTO)**

09/757100

Thus, lung tumor cDNAs are isolated and characterized using conventional cDNA library subtraction, PCR-based subtraction, microarray anal., and lung chip 5. The cDNA libraries are isolated from lung adenocarcinoma, small cell lung carcinoma, lung neuroendocrine carcinoma, and squamous cell lung carcinoma. The immunogenic gene products are expressed in Escherichia coli, and epitope mapping is provided using antibodies one of the lung tumor-specific antigens, designated L552S. The disclosed compns. are useful, for example, in the diagnosis, prevention and/or treatment of diseases, particularly lung cancer.

IT 388152-61-2P

RL: ANT (Analyte); BPN (Biosynthetic preparation); DGN (Diagnostic use); PRP (Properties); THU (Therapeutic use); ANST (Analytical study); BIOL (Biological study); PREP (Preparation); USES (Uses) (nucleotide sequence; human nucleic acid and expressed protein compns. and methods for the therapy and diagnosis of lung cancer)

L2 ANSWER 6 OF 7 HCAPLUS COPYRIGHT 2003 ACS

ACCESSION NUMBER: 2001:731174 HCAPLUS

DOCUMENT NUMBER: 135:283933

TITLE: Nucleic acid markers and methods for the measurement of sensitivity or resistance to agents used in identification, assessment, prevention, and therapy of human cancers

INVENTOR(S): Lillie, James; Brown, Jeffrey L.; Bolt, Andrew; Van Huffel, Christophe

PATENT ASSIGNEE(S): Millennium Predictive Medicine, Inc., USA

SOURCE: PCT Int. Appl., 223 pp.

CODEN: PIXXD2

DOCUMENT TYPE: Patent

LANGUAGE: English

FAMILY ACC. NUM. COUNT: 1

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2001073430	A2	20011004	WO 2001-US9202	20010322
WO 2001073430	A3	20020411		
W:	AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM			
RW:	GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG			

US 2002081596 A1 20020627 US 2001-816292 20010322

PRIORITY APPLN. INFO.: US 2000-192100P P 20000324

US 2000-197064P P 20000413

AB The present invention is directed to the identification of markers that can be used to det. whether cancer cells are sensitive or resistant to a therapeutic agent. The present invention is also directed to the identification of therapeutic targets. The invention features 436 "sensitivity markers". These are markers that are expressed in most or all cell lines that are sensitive to

**THIS PAGE BLANK (USPTO)**

09/757100

treatment with an agent and which are not expressed (or are expressed at a rather low level) in cells that are resistant to treatment with that agent. The invention also features 3132 "resistance markers". These are markers that are expressed in most or all cell lines that are resistant to treatment with an agent and which are not expressed (or are expressed at a rather low level) in cells that are sensitive to treatment with that agent. PCR-based subtractive hybridization of cDNA libraries from normal and cancerous tissues was used to identify the markers. At least some of the examples relate to sensitivity or resistance to taxol.

IT 191206-39-0

RL: ANT (Analyte); BOC (Biological occurrence); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); ANST (Analytical study); BIOL (Biological study); OCCU (Occurrence); USES (Uses)

(resistance marker; nucleic acid markers and methods for the measurement of sensitivity or resistance to agents used in identification, assessment, prevention, and therapy of human cancers)

L2 ANSWER 7 OF 7 HCAPLUS COPYRIGHT 2003 ACS

ACCESSION NUMBER: 2001:551387 HCAPLUS

DOCUMENT NUMBER: 135:103496

TITLE: Nucleic acid compositions, kits, and methods for identification, assessment, prevention, and therapy of human breast cancer

INVENTOR(S): Lillie, James; Palermo, Adam; Wang, Youzhen; Steinmann, Kathleen; Elias, Josh

PATENT ASSIGNEE(S): Millennium Predictive Medicine, Inc., USA

SOURCE: PCT Int. Appl., 2674 pp.

CODEN: PIXXD2

DOCUMENT TYPE: Patent

LANGUAGE: English

PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2001046697	A2	20010628	WO 2000-US35214	20001221
W:	AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM			
RW:	AT, BE, BF, BJ, CF, CG, CH, CI, CM, CY, DE, DK, ES, FI, FR, GA, GB, GR, IE, IT, LU, MC, ML, MR, NE, NL, PT, SE, SN, TD, TG, TR			
PRIORITY APPLN. INFO.:			US 1999-PV171406	19991221
			US 2000-PV176423	20000114
			US 2000-PV190471	20000317
			US 2000-PV193482	20000329
			US 2000-PV205231	20000515
			US 2000-PV213236	20000620
			US 2000-PV219865	20000720

AB The invention relates to nucleic acid marker compns., kits and methods for detecting, characterizing, preventing, and treating human breast cancers. A variety of markers are provided, wherein changes in the levels of expression of one or more of the nucleic acid markers is correlated with the presence of breast cancer. The

Searcher : Shears 308-4994

**THIS PAGE BLANK (USPTO)**

09/757100

level of expression of numerous potential markers was measured in cells obtained from breast cancer tissue samples obtained from fifteen patients afflicted with breast cancer and from eleven breast cancer cell cultures, based on comparison with expression levels of each marker in corresponding non-cancerous breast tissue and cell cultures. The 15 cancer tissue samples include (i) five invasive lobular carcinomas (ILC), (ii) five invasive ductal carcinomas (IDC), and (iii) five samples of ductal carcinoma in situ (DCIS). As an addnl. evaluation of ability to indicate breast cancer, individual markers that were identified by transcriptional profiling criteria were also tested in six different subtracted library expts. In addn., protein profiling expts. were undertaken to assess whether the proteins assocd. with the expression of individual markers of the invention are secreted. Table 21 lists approx. 43,500 GenBank Accession Nos. from the present invention. [This abstr. record is two of eight records for this document necessitated by the large no. of index entries required to fully index the document and publication system constraints.].

IT 191206-39-0

RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)  
(nucleotide sequence; nucleic acid compns., kits, and methods for identification, assessment, prevention, and therapy of human breast cancer)

E61 THROUGH E67 ASSIGNED

FILE 'REGISTRY' ENTERED AT 10:57:43.ON 13 JAN 2003

L3 7 SEA FILE=REGISTRY ABB=ON PLU=ON (191206-39-0/BI OR 270220-29-6/BI OR 270220-33-2/BI OR 272727-43-2/BI OR 388152-61-2/BI OR 475019-40-0/BI OR 475864-34-7/BI)

L3 ANSWER 1 OF 7 REGISTRY COPYRIGHT 2003 ACS

RN 475864-34-7 REGISTRY

CN DNA (human clone R0127:G07 lung tumor-associated protein cDNA fragment) (9CI) (CA INDEX NAME)

OTHER NAMES:

CN 981: PN: US20020172952 SEQID: 981 claimed DNA

SQL 550

MF Unspecified

CI MAN

REFERENCE 1: 137:381006

L3 ANSWER 2 OF 7 REGISTRY COPYRIGHT 2003 ACS

RN 475019-40-0 REGISTRY

CN DNA (human clone R0127:G07 lung tumor-associated protein cDNA fragment) (9CI) (CA INDEX NAME)

OTHER NAMES:

CN 981: PN: US20020168637 SEQID: 981 claimed DNA

SQL 550

MF Unspecified

CI MAN

REFERENCE 1: 137:364446

L3 ANSWER 3 OF 7 REGISTRY COPYRIGHT 2003 ACS

RN 388152-61-2 REGISTRY

Searcher : Shears 308-4994

**THIS PAGE BLANK (USPTO)**



09/757100

CN DNA (human clone B0127:G07 lung tumor-associated protein cDNA  
fragment) (9CI) (CA INDEX NAME)

OTHER NAMES:

CN 979: PN: WO0204514 SEQID: 981 claimed DNA

SQL 550

MF Unspecified

CI MAN

REFERENCE 1: 136:129944

L3 ANSWER 4 OF 7 REGISTRY COPYRIGHT 2003 ACS

RN **272727-43-2** REGISTRY

CN DNA (human clone CM2-BT0741-050400-140-c12 EST (expressed sequence  
tag)) (9CI) (CA INDEX NAME)

OTHER NAMES:

CN GenBank BE092864

SQL 412

MF Unspecified

CI MAN

REFERENCE 1: 136:351174

L3 ANSWER 5 OF 7 REGISTRY COPYRIGHT 2003 ACS

RN **270220-33-2** REGISTRY

CN DNA (human clone RC2-BN0032-120200-011-d08 EST (expressed sequence  
tag)) (9CI) (CA INDEX NAME)

OTHER NAMES:

CN GenBank AW992847

SQL 682

MF Unspecified

CI MAN

REFERENCE 1: 136:351185

L3 ANSWER 6 OF 7 REGISTRY COPYRIGHT 2003 ACS

RN **270220-29-6** REGISTRY

CN DNA (human clone RC2-BN0032-120200-011-b08 EST (expressed sequence  
tag)) (9CI) (CA INDEX NAME)

OTHER NAMES:

CN GenBank AW992843

SQL 648

MF Unspecified

CI MAN

REFERENCE 1: 136:351185

L3 ANSWER 7 OF 7 REGISTRY COPYRIGHT 2003 ACS

RN **191206-39-0** REGISTRY

CN DNA (human clone 795352 EST (expressed sequence tag)) (9CI) (CA  
INDEX NAME)

OTHER NAMES:

CN 1216: PN: WO0173430 TABLE: 2 claimed DNA

CN 4606: PN: WO0146697 TABLE: 21 claimed DNA

SQL 598

MF Unspecified

CI MAN

REFERENCE 1: 135:283933

THIS PAGE BLANK (USPTO)

09/757100

REFERENCE 2: 135:103496

FILE 'HOME' ENTERED AT 10:58:05 ON 13 JAN 2003

Searcher : Shears 308-4994

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2003, 09:46:50 ; Search time 3165 Seconds  
(without alignments)  
158.878 Million cell updates/sec

Title: US-09-757-100B-18

Perfect score: 20

Sequence: 1 tttaaccagatggtcattc 20

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 24791104 seqs, 12571243825 residues

Total number of hits satisfying chosen parameters: 49582208

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents NA Main:\*

1: /cgn2\_6/ptodata/1/pna/PCTUS COMB.seq.\*  
2: /cgn2\_6/ptodata/1/pna/US06 COMB.seq.\*  
3: /cgn2\_6/ptodata/1/pna/US07 COMB.seq.\*  
4: /cgn2\_6/ptodata/1/pna/US08 COMB.seq.\*  
5: /cgn2\_6/ptodata/1/pna/US081 COMB.seq.\*  
6: /cgn2\_6/ptodata/1/pna/US082 COMB.seq.\*  
7: /cgn2\_6/ptodata/1/pna/US083 COMB.seq.\*  
8: /cgn2\_6/ptodata/1/pna/US084 COMB.seq.\*  
9: /cgn2\_6/ptodata/1/pna/US085 COMB.seq.\*  
10: /cgn2\_6/ptodata/1/pna/US086 COMB.seq.\*  
11: /cgn2\_6/ptodata/1/pna/US087 COMB.seq.\*  
12: /cgn2\_6/ptodata/1/pna/US088 COMB.seq.\*  
13: /cgn2\_6/ptodata/1/pna/US089 COMB.seq.\*  
14: /cgn2\_6/ptodata/1/pna/US090 COMB.seq.\*  
15: /cgn2\_6/ptodata/1/pna/US091 COMB.seq.\*  
16: /cgn2\_6/ptodata/1/pna/US092 COMB.seq.\*  
17: /cgn2\_6/ptodata/1/pna/US093 COMB.seq.\*  
18: /cgn2\_6/ptodata/1/pna/US094 COMB.seq.\*  
19: /cgn2\_6/ptodata/1/pna/US095A COMB.seq.\*  
20: /cgn2\_6/ptodata/1/pna/US095B COMB.seq.\*  
21: /cgn2\_6/ptodata/1/pna/US095C COMB.seq.\*  
22: /cgn2\_6/ptodata/1/pna/US095D COMB.seq.\*  
23: /cgn2\_6/ptodata/1/pna/US096A COMB.seq.\*  
24: /cgn2\_6/ptodata/1/pna/US096B COMB.seq.\*  
25: /cgn2\_6/ptodata/1/pna/US096C COMB.seq.\*  
26: /cgn2\_6/ptodata/1/pna/US096D COMB.seq.\*  
27: /cgn2\_6/ptodata/1/pna/US096E COMB.seq.\*  
28: /cgn2\_6/ptodata/1/pna/US097A COMB.seq.\*  
29: /cgn2\_6/ptodata/1/pna/US097B COMB.seq.\*  
30: /cgn2\_6/ptodata/1/pna/US097C COMB.seq.\*  
31: /cgn2\_6/ptodata/1/pna/US098A COMB.seq.\*  
32: /cgn2\_6/ptodata/1/pna/US098B COMB.seq.\*  
33: /cgn2\_6/ptodata/1/pna/US098C COMB.seq.\*  
34: /cgn2\_6/ptodata/1/pna/US099A COMB.seq.\*  
35: /cgn2\_6/ptodata/1/pna/US099B COMB.seq.\*  
36: /cgn2\_6/ptodata/1/pna/US099C COMB.seq.\*  
37: /cgn2\_6/ptodata/1/pna/US099D COMB.seq.\*  
38: /cgn2\_6/ptodata/1/pna/US100A COMB.seq.\*  
39: /cgn2\_6/ptodata/1/pna/US100B COMB.seq.\*  
40: /cgn2\_6/ptodata/1/pna/US101A COMB.seq.\*  
41: /cgn2\_6/ptodata/1/pna/US101B COMB.seq.\*  
42: /cgn2\_6/ptodata/1/pna/US102A COMB.seq.\*  
43: /cgn2\_6/ptodata/1/pna/US102B COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	1	PCT-US00-18999-18
2	20	100.0	20	17	US-09-377-310-18
3	20	100.0	20	29	US-09-757-100B-18
C 4	20	100.0	325	17	US-09-332-782-6610
C 5	20	100.0	325	19	US-09-515-694-6610
C 6	20	100.0	379	17	US-09-359-067-23180
C 7	20	100.0	416	38	US-10-007-926A-284
C 8	20	100.0	421	31	US-09-824-557-1277
C 9	20	100.0	433	30	US-09-790-435-4313
C 10	20	100.0	436	16	US-09-289-768-4292
C 11	20	100.0	436	35	US-09-939-397-4292
C 12	20	100.0	443	18	US-09-431-517-3422
C 13	20	100.0	449	17	US-09-306-350A-30669
C 14	20	100.0	449	34	US-09-909-629-30669
C 15	20	100.0	455	31	US-09-824-557-3555
C 16	20	100.0	456	20	US-09-532-315-27166
C 17	20	100.0	470	16	US-09-287-618-24792
C 18	20	100.0	542	17	US-09-371-188-3928
C 19	20	100.0	542	31	US-09-801-833-3928
C 20	20	100.0	550	23	US-09-614-124-981
21	20	100.0	550	23	US-09-614-124B-981

Lacourciere  
091757102  
Seq. ID 18 w  
Interf

```
22 20 100.0 550 25 US-09-651-563-981 Sequence 981, App
23 20 100.0 550 25 US-08-658-824-981 Sequence 981, App
24 20 100.0 550 26 US-09-671-325-981 Sequence 981, App
25 20 100.0 550 26 US-09-677-419A-981 Sequence 981, App
26 20 100.0 550 28 US-09-702-705-981 Sequence 981, App
27 20 100.0 550 29 US-09-736-457-981 Sequence 981, App
28 20 100.0 550 32 US-09-849-626-981 Sequence 981, App
29 20 100.0 550 34 US-09-902-941-981 Sequence 981, App
30 20 100.0 550 38 US-10-017-754-981 Sequence 981, App
31 20 100.0 550 40 US-10-113-872-981 Sequence 981, App
32 20 100.0 551 31 US-09-823-327-7081 Sequence 7081, App
33 20 100.0 751 17 US-09-396-885-5123 Sequence 5123, App
34 20 100.0 751 17 US-09-396-885A-5123 Sequence 5123, App
35 20 100.0 1555 71 US-60-278-358-3581 Sequence 3581, App
36 20 100.0 1561 29 US-09-758-467-159 Sequence 159, App
37 20 100.0 1561 42 US-10-238-810-159 Sequence 159, App
38 20 100.0 1566 67 US-60-234-690-6983 Sequence 6983, App
39 20 100.0 1566 76 US-60-333-843-6983 Sequence 6983, App
40 20 100.0 1803 29 US-09-758-467-62 Sequence 62, App
41 20 100.0 1803 42 US-10-238-810-62 Sequence 62, App
42 20 100.0 3052 38 US-10-007-926A-285 Sequence 285, App
43 20 100.0 3598 17 US-09-359-922-4921 Sequence 4921, App
44 20 100.0 3598 17 US-09-359-922-4921 Sequence 4921, App
45 20 100.0 3598 34 US-09-919-002-4921 Sequence 4921, App
```

## ALIGNMENTS

```
RESULT 1
PCT-US00-18999-18
; Sequence 18, Application PC/TUS0018999
; GENERAL INFORMATION:
; APPLICANT: Isis Pharmaceuticals, Inc.
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; APPLICANT: Nero, Pamela S.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; TITLE OF INVENTION: Expression
; FILE REFERENCE: ISPH-0476
; CURRENT APPLICATION NUMBER: PCT/US00/18999
; CURRENT FILING DATE: 2000-07-13
; PRIOR APPLICATION NUMBER: 09/377,310
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
PCT-US00-18999-18

Query Match 100.0%; Score 20; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TTTCAACCGAGTGTGTCATTC 20
Db 1 TTTCAACCGAGTGTGTCATTC 20

RESULT 2
US-09-377-310-18
; Sequence 18, Application US/09377310A
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; TITLE OF INVENTION: Expression
; FILE REFERENCE: ISPH-0389
; CURRENT APPLICATION NUMBER: US/09/377,310A
; CURRENT FILING DATE: 1999-08-19
```

```
; CURRENT FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-377-310-18
```

```
Query Match 100.0%; Score 20; DB 17; Length 20;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 1 TTTCAACCGAGTGTGTCATTC 20
Db 1 TTTCAACCGAGTGTGTCATTC 20
```

```
RESULT 3
US-09-757-100B-18
; Sequence 18, Application US/09757100B
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; APPLICANT: Nero, Pamela S.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; TITLE OF INVENTION: Expression
; FILE REFERENCE: ISPH-0533
; CURRENT APPLICATION NUMBER: US/09/757,100B
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 09/377,310
; PRIOR FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: PCT/US00/18999
; PRIOR FILING DATE: 2000-07-13
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-757-100B-18
```

```
Query Match 100.0%; Score 20; DB 29; Length 20;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 1 TTTCAACCGAGTGTGTCATTC 20
Db 1 TTTCAACCGAGTGTGTCATTC 20
```

```
RESULT 4
US-09-332-782-6610/c
; Sequence 6610, Application US/09332782
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-754CON1
; CURRENT APPLICATION NUMBER: US/09/332,782
; CURRENT FILING DATE: 1999-06-14
; EARLIER APPLICATION NUMBER: US 09/181,430
; EARLIER FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 21027
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6610
; LENGTH: 325
; TYPE: DNA
; ORGANISM: Homo sapiens
```

US-09-332-782-6610

Query Match 100.0%; Score 20; DB 17; Length 325;  
Best Local Similarity 100.0%; Pred. No. 9.5;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTTCACACGAGTGGTCATTC 20  
|||||  
DB 53 TTTCACACGAGTGGTCATTC 34

## RESULT 5

US-09-515-694-6610/c  
; Sequence 6610, Application US/09515694  
; GENERAL INFORMATION:

; APPLICANT: Arterburn, Matthew  
; APPLICANT: Asghari, Vida  
; APPLICANT: Blun, Linda  
; APPLICANT: Cheung, Patrick  
; APPLICANT: Damavandi, Simin  
; APPLICANT: Dickson, Mark  
; APPLICANT: Drake, Jim  
; APPLICANT: Drmanac, Radoje  
; APPLICANT: Engleman, Carrie  
; APPLICANT: Faulkner, Brandy  
; APPLICANT: Fox, Melvin  
; APPLICANT: Garcia, Veronica  
; APPLICANT: Giedt, Gretchen  
; APPLICANT: Jomek, Leni  
; APPLICANT: Jones, Lee  
; APPLICANT: Kita, David  
; APPLICANT: Kofler, Janette  
; APPLICANT: Labat, Ivan  
; APPLICANT: Lee, Won-Jae  
; APPLICANT: Lomelli, Michelle  
; APPLICANT: Nguyen, Hong  
; APPLICANT: Nguyen, Linh  
; APPLICANT: Nguyen, Lynne  
; APPLICANT: Nguyen, Phuong  
; APPLICANT: Nogra, Margie  
; APPLICANT: Ojeda, Jesse  
; APPLICANT: Palencia, Servando  
; APPLICANT: Rasi, Fariba  
; APPLICANT: Randhwa, Gurpreet  
; APPLICANT: Sanourieh, Hannah  
; APPLICANT: Sidhu, Navjwan  
; APPLICANT: Smith, Benjamin  
; APPLICANT: Smythe, Ashleigh  
; APPLICANT: Tkach, Joe  
; APPLICANT: Tulpule, Mukul  
; APPLICANT: Verna, Ron  
; APPLICANT: Wachter, Adam  
; APPLICANT: Wu, James  
; APPLICANT: Yim, Kenneth  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
; FILE REFERENCE: 754CIP  
; CURRENT APPLICATION NUMBER: US/09/515,694  
; CURRENT FILING DATE: 2000-02-29  
; EARLIER APPLICATION NUMBER: 09/332,782  
; EARLIER FILING DATE: 1999-06-14  
; EARLIER APPLICATION NUMBER: 09/181,430  
; EARLIER FILING DATE: 1998-10-28  
; NUMBER OF SEQ ID NOS: 21027  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 6610  
; LENGTH: 325  
; TYPE: DNA  
; ORGANISM: Homo sapiens

US-09-515-694-6610

Query Match 100.0%; Score 20; DB 19; Length 325;  
Best Local Similarity 100.0%; Pred. No. 9.5;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTTCACACGAGTGGTCATTC 20  
|||||  
DB 53 TTTCACACGAGTGGTCATTC 34

## RESULT 6

US-09-359-067-23180/c  
; Sequence 23180, Application US/09359067  
; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc.  
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
; FILE REFERENCE: 20411-747CON1  
; CURRENT APPLICATION NUMBER: US/09/359,067  
; CURRENT FILING DATE: 1999-07-22  
; EARLIER APPLICATION NUMBER: US 09/131,598  
; EARLIER FILING DATE: 1998-08-10  
; NUMBER OF SEQ ID NOS: 49786  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 23180  
; LENGTH: 379  
; TYPE: DNA  
; ORGANISM: Homo sapiens

US-09-359-067-23180

Query Match 100.0%; Score 20; DB 17; Length 379;  
Best Local Similarity 100.0%; Pred. No. 9.8;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTTCACACGAGTGGTCATTC 20  
|||||  
DB 339 TTTCACACGAGTGGTCATTC 320

## RESULT 7

US-10-007-926A-284/c  
; Sequence 284, Application US/10007926A  
; GENERAL INFORMATION:

; APPLICANT: BERTUCCI, FRANCOIS  
; APPLICANT: HOULGATE, REMI  
; APPLICANT: BIRNBAUM, DANIEL  
; APPLICANT: NGUYEN, CATHERINE  
; APPLICANT: VIENS, PATRICE  
; APPLICANT: FERT, VINCENT  
; TITLE OF INVENTION: GENE EXPRESSION PROFILING OF PRIMARY BREAST CARCINOMAS  
; FILE REFERENCE: 1546-R-00  
; CURRENT APPLICATION NUMBER: US/10/007,926A  
; CURRENT FILING DATE: 2001-12-07  
; PRIOR APPLICATION NUMBER: 60/254,090  
; PRIOR FILING DATE: 2000-12-08  
; NUMBER OF SEQ ID NOS: 468  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 284  
; LENGTH: 416  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: 5' terminal sequence. ptk2 protein tyrosine  
; OTHER INFORMATION: kinase 2 (PTK2) gene.  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (275)  
; OTHER INFORMATION: a, t, c or g  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (295)  
; OTHER INFORMATION: a, t, c or g  
; FEATURE:  
; NAME/KEY: modified\_base  
; LOCATION: (364)

OTHER INFORMATION: a, t, c or g  
US-10-007-926A-284

Query Match 100.0%; Score 20; DB 38; Length 416;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTTCACCGAGATGTCATTC 20  
|||||  
DB 343 TTTCACCGAGATGTCATTC 324

RESULT 8  
US-09-824-557-1277/C  
Sequence 1277, Application US/09824557  
GENERAL INFORMATION:  
APPLICANT: Shyjan, Andrew W.  
TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES  
TITLE OF INVENTION: THEREFOR  
FILE REFERENCE: 1600.2086-001  
CURRENT APPLICATION NUMBER: US/09/824,557  
CURRENT FILING DATE: 2001-04-02  
PRIOR APPLICATION NUMBER: US 60/193,438  
PRIOR FILING DATE: 2000-03-31  
NUMBER OF SEQ ID NOS: 9773  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1277  
LENGTH: 421  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(421)  
OTHER INFORMATION: n = A,T,C or G  
US-09-824-557-1277

Query Match 100.0%; Score 20; DB 31; Length 421;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTTCACCGAGATGTCATTC 20  
|||||  
DB 270 TTTCACCGAGATGTCATTC 251

RESULT 9  
US-09-790-435-4313/C  
Sequence 4313, Application US/09790435  
GENERAL INFORMATION:  
APPLICANT: Holtzman, Douglas A.  
APPLICANT: Geating, David P.  
TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES  
TITLE OF INVENTION: THEREFOR  
FILE REFERENCE: 1600.2060-001  
CURRENT APPLICATION NUMBER: US/09/790,435  
CURRENT FILING DATE: 2001-02-22  
PRIOR APPLICATION NUMBER: US 60/183,730  
PRIOR FILING DATE: 2000-02-22  
NUMBER OF SEQ ID NOS: 7473  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4313  
LENGTH: 433  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(433)  
OTHER INFORMATION: n = A,T,C or G  
US-09-790-435-4313

Query Match 100.0%; Score 20; DB 30; Length 433;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTTCACCGAGATGTCATTC 20  
|||||  
DB 200 TTTCACCGAGATGTCATTC 181

RESULT 10  
US-09-289-768-4292/C  
Sequence 4292, Application US/09289768  
GENERAL INFORMATION:  
APPLICANT: Hyseq, Inc.  
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES  
FILE REFERENCE: 20411-765  
CURRENT APPLICATION NUMBER: US/09/289,768  
CURRENT FILING DATE: 1999-04-08  
NUMBER OF SEQ ID NOS: 39996  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 4292  
LENGTH: 436  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(436)  
OTHER INFORMATION: n = A,T,C or G  
US-09-289-768-4292

Query Match 100.0%; Score 20; DB 16; Length 436;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTTCACCGAGATGTCATTC 20  
|||||  
DB 97 TTTCACCGAGATGTCATTC 78

RESULT 11  
US-09-939-397-4292/C  
Sequence 4292, Application US/09939397  
GENERAL INFORMATION:  
APPLICANT: Hyseq, Inc.  
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED  
TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES  
FILE REFERENCE: 20411-765  
CURRENT APPLICATION NUMBER: US/09/939,397  
CURRENT FILING DATE: 2001-08-24  
PRIOR APPLICATION NUMBER: 09/289,768  
PRIOR FILING DATE: 1999-04-08  
NUMBER OF SEQ ID NOS: 39996  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 4292  
LENGTH: 436  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(436)  
OTHER INFORMATION: n = A,T,C or G  
US-09-939-397-4292

Query Match 100.0%; Score 20; DB 35; Length 436;  
Best Local Similarity 100.0%; Pred. No. 10;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTTCACCGAGATGTCATTC 20  
|||||  
DB 97 TTTCACCGAGATGTCATTC 78

RESULT 12  
US-09-431-517-3422/C  
Sequence 3422, Application US/09431517



```
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-751CON1
; CURRENT APPLICATION NUMBER: US/09/431,517
; CURRENT FILING DATE: 1999-11-01
; EARLIER APPLICATION NUMBER: US 09/170,294
; EARLIER FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 31760
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3422
; LENGTH: 443
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(443)
; OTHER INFORMATION: n = A,T,C or G
US-09-431-517-3422
```

```
Query Match 100.0%; Score 20; DB 18; Length 443;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 TTTCACACAGATGGTCATTC 20
Db 73 TTTCACACAGATGGTCATTC 54
```

```
RESULT 13
US-09-306-350A-30669/c
; Sequence 30669, Application US/09306350A
; GENERAL INFORMATION:
; APPLICANT: Dickson, Mark C.
; APPLICANT: Drmanac, Radoje T.
; APPLICANT: Jones, Lee W.
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; FILE REFERENCE: 20411-776
; CURRENT APPLICATION NUMBER: US/09/306,350A
; CURRENT FILING DATE: 1999-05-07
; NUMBER OF SEQ ID NOS: 41304
; SOFTWARE: pt CT_1 Version 1.1
; SEQ ID NO 30669
; LENGTH: 449
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-306-350A-30669
```

```
Query Match 100.0%; Score 20; DB 17; Length 449;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 TTTCACACAGATGGTCATTC 20
Db 208 TTTCACACAGATGGTCATTC 189
```

```
RESULT 14
US-09-629-30669/c
; Sequence 30669, Application US/09909629
; GENERAL INFORMATION:
; APPLICANT: Dickson, Mark C.
; APPLICANT: Drmanac, Radoje T.
; APPLICANT: Jones, Lee W.
; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; FILE REFERENCE: 20411-776
```

```
; CURRENT APPLICATION NUMBER: US/09/909,629
; CURRENT FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: 09/306,350
; PRIOR FILING DATE: 1999-05-07
; NUMBER OF SEQ ID NOS: 41304
; SOFTWARE: pt CT_1 Version 1.1
; SEQ ID NO 30669
; LENGTH: 449
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-909-629-30669
```

```
Query Match 100.0%; Score 20; DB 34; Length 449;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 TTTCACACAGATGGTCATTC 20
Db 208 TTTCACACAGATGGTCATTC 189
```

```
RESULT 15
US-09-824-557-3555/c
; Sequence 3555, Application US/09824557
; GENERAL INFORMATION:
; APPLICANT: Stryjan, Andrew W.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.2086-001
; CURRENT APPLICATION NUMBER: US/09/824,557
; CURRENT FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: US 60/193,438
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 9773
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3555
; LENGTH: 455
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(455)
; OTHER INFORMATION: n = A,T,C or G
US-09-824-557-3555
```

```
Query Match 100.0%; Score 20; DB 31; Length 455;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 TTTCACACAGATGGTCATTC 20
Db 266 TTTCACACAGATGGTCATTC 247
```

```
Search completed: January 12, 2003, 12:04:06
Job time : 3167 secs
```



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2003, 09:51:30 ; Search time 168 Seconds  
(without alignments)  
110.212 Million cell updates/sec

Title: US-09-757-100B-18

Perfect score: 20  
Sequence: 1 ttccaaccagatggtcattc 20

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2355679 seqs, 462891617 residues

Total number of hits satisfying chosen parameters: 4711358

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending Patents NA New.\*

- 1: /cgn2\_6/prodata/1/pna/PCT\_NEW\_COMB.seq.\*
- 2: /cgn2\_6/prodata/1/pna/US06\_NEW\_COMB.seq.\*
- 3: /cgn2\_6/prodata/1/pna/US07\_NEW\_COMB.seq.\*
- 4: /cgn2\_6/prodata/1/pna/US08\_NEW\_COMB.seq.\*
- 5: /cgn2\_6/prodata/1/pna/US09\_NEW\_COMB.seq.\*
- 6: /cgn2\_6/prodata/1/pna/US10\_NEW\_COMB.seq.\*
- 7: /cgn2\_6/prodata/1/pna/US60\_NEW\_COMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	550	1	PCT-US02-34777-981
2	20	100.0	550	6	US-10-283-017-981
3	20	100.0	1317	5	US-09-724-676-20756
4	20	100.0	1317	5	US-09-724-676-20756
5	20	100.0	1362	5	US-09-724-676-20755
6	20	100.0	1362	5	US-09-724-676-20755
7	20	100.0	1791	5	US-09-724-676-20750
8	20	100.0	1791	5	US-09-724-676-20750
9	20	100.0	1836	5	US-09-724-676-20749
10	20	100.0	1836	5	US-09-724-676-20749
11	20	100.0	3610	5	US-09-724-676-20752
12	20	100.0	3610	5	US-09-724-676-20752
13	20	100.0	3677	5	US-09-724-676-20751
14	20	100.0	3677	5	US-09-724-676-20751
15	20	100.0	3697	5	US-09-724-676-20754
16	20	100.0	3697	5	US-09-724-676-20754
17	20	100.0	3764	5	US-09-724-676-20753
18	20	100.0	3764	5	US-09-724-676-20753
19	20	100.0	4371	5	US-09-724-676-20759
20	20	100.0	4371	5	US-09-724-676-20759
21	20	100.0	4438	5	US-09-724-676-20758
22	20	100.0	4438	5	US-09-724-676-20758
23	20	100.0	4458	5	US-09-724-676-20748
24	20	100.0	4458	5	US-09-724-676-20748
25	20	100.0	4481	5	US-10-240-965-92
26	20	100.0	4525	5	US-09-724-676-20760

ALIGNMENTS

RESULT 1

PCT-US02-34777-981

; Sequence 981, Application PC/TUS0234777

; GENERAL INFORMATION:

; APPLICANT: Corixa Corporation

; APPLICANT: Henderson, Robert A.

; APPLICANT: Wang, Tongtong

; APPLICANT: Watanabe, Yoshihiro

; APPLICANT: Kalos, Michael D.

; APPLICANT: Sleath, Paul R.

; APPLICANT: Johnson, Jeffrey C.

; APPLICANT: Retter, Marc W.

; APPLICANT: Durham, Margarita

; APPLICANT: Carter, Darick

; APPLICANT: Fanger, Gary R.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: McNabb, Andria

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE REFERENCE: 210121.47803PC

; CURRENT APPLICATION NUMBER: PCT/US02/34777

; CURRENT FILING DATE: 2002-10-28

; NUMBER OF SEQ ID NOS: 2157

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 981

; LENGTH: 550

; TYPE: DNA

; ORGANISM: Homo sapiens

PCT-US02-34777-981

Query Match 100.0%; Score 20; DB 1; Length 550;

Best Local Similarity 100.0%; Pred. No. 0.47;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTTCACACAGATGTCATTC 20

|||||

Db 195 TTTCACACAGATGTCATTC 214

RESULT 2

US-10-283-017-981

; Sequence 981, Application US/10283017

; GENERAL INFORMATION:

; APPLICANT: Henderson, Robert A.

; APPLICANT: Wang, Tongtong

; APPLICANT: Watanabe, Yoshihiro

; APPLICANT: Kalos, Michael D.

; APPLICANT: Sleath, Paul R.

```

; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Durham, Margalita
; APPLICANT: Carter, Darick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.478C20
; CURRENT APPLICATION NUMBER: US/10/283,017
; CURRENT FILING DATE: 2002-10-28
; NUMBER OF SEQ ID NOS: 2157
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 981
; LENGTH: 550
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-283-017-981
```

```
Query Match          100.0%; Score 20; DB 6; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.47;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 1 TTTCAACCGAGTGTGTCATTC 20
    |||||
Db 195 TTTCAACCGAGTGTGTCATTC 214
```

```

RESULT 3
US-09-724-676-20756/c
; Sequence 20756, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20756
; LENGTH: 1317
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-20756
```

```
Query Match          100.0%; Score 20; DB 5; Length 1317;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 1 TTTCAACCGAGTGTGTCATTC 20
    |||||
Db 372 TTTCAACCGAGTGTGTCATTC 353
```

```

RESULT 4
US-09-724-676A-20756/c
; Sequence 20756, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20756
; LENGTH: 1317
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-20756
```

```
Query Match          100.0%; Score 20; DB 5; Length 1317;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 1 TTTCAACCGAGTGTGTCATTC 20
    |||||
Db 372 TTTCAACCGAGTGTGTCATTC 353
```

```

RESULT 5
US-09-724-676-20755/c
; Sequence 20755, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20755
; LENGTH: 1362
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-20755
```

```
Query Match          100.0%; Score 20; DB 5; Length 1362;
Best Local Similarity 100.0%; Pred. No. 0.54;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 1 TTTCAACCGAGTGTGTCATTC 20
    |||||
Db 417 TTTCAACCGAGTGTGTCATTC 398
```

```

RESULT 6
US-09-724-676A-20755/c
; Sequence 20755, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20755
; LENGTH: 1362
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-20755
```

```
Query Match          100.0%; Score 20; DB 5; Length 1362;
Best Local Similarity 100.0%; Pred. No. 0.54;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy 1 TTTCAACCGAGTGTGTCATTC 20
    |||||
Db 417 TTTCAACCGAGTGTGTCATTC 398
```

```

RESULT 7
US-09-724-676-20750/c
; Sequence 20750, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20750
```

```

; LENGTH: 1791
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-20750

```

Query Match 100.0%; Score 20; DB 5; Length 1791;  
Best Local Similarity 100.0%; Pred. No. 0.56;  
Matches 20; Conservative 0; Mismatches 0; Indels

QY 1 TTTCAACCAGATGGTCATTC 20  
Db 846 TTTCAACCAGATGGTCATTC 827

RESULT 8  
US-09-724-676A-20750/c  
; Sequence 20750, Application US/09724676A

```

; ORGANISM:
; APPLICANT: Comugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Comugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20750
; LENGTH: 1791
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-20750

```

Query Match 100.0%; Score 20; DB 5; Length 1791;  
Best Local Similarity 100.0%; Pred. No. 0.56;  
Matches 20; Conservative 0; Mismatches 0; Indels

1 TTTCAACCAGATGGTCATTC 20  
|||  
846 TTTCAACCAGATGGTCATTC 827

RESULT 9  
US-09-724-676-20749/c  
; Sequence 20749, Application US/09724676

```

; ORGANISM: CompuGen LTD
; APPLICANT: CompuGen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 CompuGen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20749
; LENGTH: 1836
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-20749

```

Query Match 100.0%; Score 20; DB 5; Length 1836;  
Best Local Similarity 100.0%; Pred. No. 0.56;  
Matches 20; Conservative 0; Mismatches 0; Indels

Qy 1 TTTCAACCAGATGGTCATTC 20  
Db 891 TTTCAACCAGATGGTCATTC 872

RESULT 10  
US-09-724-676A-20749/c  
; Sequence 20749. Application US/09724676A

```

; GENERAL INFORMATION:
; APPLICANT: Comugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Comugen

```

```

; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2

```

SEQ ID NO 20749  
LENGTH: 1836  
TYPE: DNA  
ORGANISM: Homo sapiens

US-09-124-676A-20149	Query Match	100.0%
	Best Local Similarity	100.0%

```

Qy      1  TTTCAACCAAGTGGTCATTC 20
      |||||
Db      891  TTTCAACCAAGTGGTCATTC 872

RESULT 11
US-09-724-676-20752/C
; Sequence 20752, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Comugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Comugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20752
; LENGTH: 3610
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-20752

```

```
Query Match      100.0%; Score 20; DB 5; Length 3610;
Best Local Similarity 100.0%; Pred. No. 0.62;
Matches 20; Conservative 0; Mismatches 0; Indels
```

QY 1 TTTCAACCAAGATGGTCATTC 20  
Db 2665 TTTCAACCAAGATGGTCATTC 2646

```

RESULT 12
US-09-724-676A-20752/c
; Sequence 20752, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: CompuGen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 CompuGen
; CURRENT APPLICATION NUMBER: US/09/724, 676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20752
; LENGTH: 3610
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-20752

```

```
Query Match      100.0%; Score 20; DB 5; Length 3610;
Best Local Similarity 100.0%; Pred.No. 0.62;
Matches 20; Conservative 0; Mismatches 0; Indels
```

Qy 1 TTTCAACCAGATGGTCATTC 20  
 |||||  
 Db 2665 TTTCAACCAGATGGTCATTC 2646

RESULT 13  
US-09-724-676-20751/c

```

; Sequence 20751, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20751
; LENGTH: 3677
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-20751

```

```

Query Match          100.0%; Score 20; DB 5; Length 3677;
Best Local Similarity 100.0%; Pred. No. 0.62;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTTCAACGAGATGTCATTC 20
Db 2732 TTTCAACGAGATGTCATTC 2713

```

```

RESULT 14
US-09-724-676A-20751/C
; Sequence 20751, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20751
; LENGTH: 3677
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-20751

```

```

Query Match          100.0%; Score 20; DB 5; Length 3677;
Best Local Similarity 100.0%; Pred. No. 0.62;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTTCAACGAGATGTCATTC 20
Db 2732 TTTCAACGAGATGTCATTC 2713

```

```

RESULT 15
US-09-724-676-20754/C
; Sequence 20754, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 20754
; LENGTH: 3697
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-20754

```

```

Query Match          100.0%; Score 20; DB 5; Length 3697;
Best Local Similarity 100.0%; Pred. No. 0.62;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTTCAACGAGATGTCATTC 20

```

Db 2752 TTTCAACGAGATGTCATTC 2733

Search completed: January 12, 2003, 12:07:16  
Job time : 170 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: January 12, 2003, 09:41:35 ; Search time 2677 Seconds  
(without alignments)  
217.429 Million cell updates/sec

Title: US-09-757-100B-18  
Perfect score: 20  
Sequence: 1 ttccaaccagatggtcattc 20

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_hgt.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sta.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mur.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sta.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_hgt\_hum.\*
- 31: em\_hgt\_inv.\*
- 32: em\_hgt\_other.\*
- 33: em\_hgt\_mus.\*
- 34: em\_hgt\_pln.\*
- 35: em\_hgt\_rod.\*
- 36: em\_hgt\_mam.\*
- 37: em\_hgt\_vrt.\*
- 38: em\_sy.\*
- 39: em\_hgtgo\_hum.\*
- 40: em\_hgtgo\_mus.\*
- 41: em\_hgtgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	20	100.0	20	6	AR116330	AR116330 Sequence
2	20	100.0	402	11	G06663	G06663 human STS W
3	20	100.0	550	6	AX368271	AX368271 Sequence
4	20	100.0	3052	9	HUMFAK	L05186 Homo sapien
5	20	100.0	3199	9	BC028733	BC028733 Homo sapi
6	20	100.0	3236	9	AK094999	AK094999 Homo sapi
7	20	100.0	3791	6	AR116314	AR116314 Sequence
8	20	100.0	3791	9	HSMFAKX	L13616 Human focal
9	20	100.0	4431	9	HSM804272	AL832961 Homo sapi
10	20	100.0	4481	6	AX281683	AX281683 Sequence
11	20	100.0	206467	2	AC067931	AC067931 Homo sapi
12	18.4	92.0	3974	1	RME132004	AJ132004 Rhizobium
13	18.4	92.0	323450	1	SME591790	AL591790 Sinorhizo
14	18	90.0	119671	9	AC102801	AC102801 Homo sapi
15	18	90.0	156235	9	AC084336	AC084336 Homo sapi
16	18	90.0	175743	9	AC011632	AC011632 Homo sapi
17	17.4	87.0	10731	1	AE005765	AE005765 Caulobact
18	17.4	87.0	50711	2	AC103694	AC103694 Homo sapi
19	17.4	87.0	102194	10	AL713916	AL713916 Mouse DNA
20	17.4	87.0	113332	2	AC091767	AC091767 Homo sapi
21	17.4	87.0	150206	8	OSJN00083	AL606668 Oryza sat
22	17.4	87.0	155902	9	AP002852	AP002852 Homo sapi
23	17.4	87.0	182551	2	AC108406	AC108406 Mus muscu
24	17.4	87.0	184332	2	AC011777	AC011777 Homo sapi
25	17.4	87.0	187918	2	AL808135	AL808135 Mus muscu
26	17.4	87.0	188895	2	AC106524	AC106524 Rattus no
27	17.4	87.0	196766	2	AC094333	AC094333 Rattus no
28	17.4	87.0	197294	2	AC111560	AC111560 Rattus no
29	17.4	87.0	200610	9	AP002851	AP002851 Homo sapi
30	17.4	87.0	208747	2	AC105614	AC105614 Rattus no
31	17.4	87.0	246700	2	AL808136	AL808136 Mus muscu
32	17.4	87.0	249536	10	AL663072	AL663072 Mouse DNA
33	17.4	87.0	275548	2	AC098455	AC098455 Rattus no
34	17	85.0	1310	8	AB070838	AB070838 Coriolus
35	17	85.0	78782	2	AC095594	AC095594 Rattus no
36	16.8	84.0	1333	4	MEBCASMR	X54715 M.eugenii m
37	16.8	84.0	1911	6	AX481431	AX481431 Sequence
38	16.8	84.0	4465	6	AR007225	AR007225 Sequence
39	16.8	84.0	4465	6	AR062429	AR062429 Sequence
40	16.8	84.0	4465	6	AR170011	AR170011 Sequence
41	16.8	84.0	4465	6	AR170438	AR170438 Sequence
42	16.8	84.0	4544	6	I27715	I27715 Sequence 1
43	16.8	84.0	4544	6	I06337	I06337 Sequence 20
44	16.8	84.0	5427	6	AR036470	AR036470 Sequence
45	16.8	84.0	5427	6	AR069429	AR069429 Sequence

ALIGNMENTS

RESULT 1  
AR116330  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES

AR116330  
Sequence 18 from patent US 6133031.  
AR116330  
AR116330.1 GI:14096652  
Unknown.  
Unknown.  
Unclassified.  
1 (bases 1 to 20)  
Monia,B.P. and Gaarde,W.A.  
Antisense inhibition of focal adhesion kinase expression  
Patent: US 6133031-A 18 17-OCT-2000;  
Location/Qualifiers

20 bp  
DNA  
linear  
PAT 16-MAY-2001

source 1. .20  
/organism="unknown"

BASE COUNT 5 a 5 c 3 g 7 t

ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 20;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTTCAACGAGATGTCATTC 20  
1 TTTCAACGAGATGTCATTC 20

Db 1 TTTCAACGAGATGTCATTC 20

RESULT 2  
G06663/c 402 bp DNA linear STS 19-OCT-1995

LOCUS G06663 human STS WI-7711, sequence tagged site.

DEFINITION G06663

ACCESSION G06663.1 GI:859908

VERSION STS; STS sequence; primer; sequence tagged site.

KEYWORDS Homo sapiens STS derived from sequences in dbEST and the Unigene collection.

SOURCE

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 402)  
Hudson,T.  
Whitehead Institute/MIT Center for Genome Research; Physically  
Mapped ESTs  
Unpublished (1995)

JOURNAL

COMMENT  
Contact: Thomas Hudson  
Whitehead Institute/MIT Center for Genome Research  
Whitehead Institute for Biomedical Research  
9 Cambridge Center, Cambridge MA 02142 USA  
Tel: 617 252 1900  
Fax: 617 252 1902  
Email: thudson@genome.wi.mit.edu

Primer A: GAGCAGCTGCTACCCCTC  
Primer B: AGTTAGGAATTAGATGTCACCC  
STS size: 350  
PCR Profile:  
Presoak:  
Denaturation:  
Annealing: 56 degrees C  
Polymerization:  
PCR Cycles: 35  
Thermal Cycler:  
Protocol:  
Template: 10 ng  
Primer: each 5 pm  
dNTPs: each 4 mM  
Tag Polymerase: 0.025 units/ul  
Total Vol: 20 ul

Buffer:  
MgCl2: 1.5 mM  
KCl: 50 mM  
Tris-HCL: 10 mM  
pH: 9.3

FEATURES  
source  
1. .402  
Location/Qualifiers  
/Organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="951\_H\_5; 805.4 CR from top of Chr8 linkage group"

STS  
primer\_bind 14. .363  
complement(339..363)

BASE COUNT 117 a 81 c 74 g 111 t 19 others

ORIGIN

Query Match 100.0%; Score 20; DB 11; Length 402;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTTCAACGAGATGTCATTC 20  
140 TTTCAACGAGATGTCATTC 121

Db 140 TTTCAACGAGATGTCATTC 121

RESULT 3  
AX368271 550 bp DNA linear PAT 16-FEB-2002

LOCUS AX368271

DEFINITION Sequence 981 from Patent WO0204514.

ACCESSION AX368271

VERSION AX368271.1 GI:18856344

KEYWORDS

SOURCE

ORGANISM human.  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1  
Wang,T., Maranabe,Y., Henderson,R.A., Johnson,J.C., Retter,M.W.,  
Marnerakis,M., Carter,D., Fanger,G.R., Veddick,T.S., Bangur,C.S.,  
McNabb,A., Fanger,N., Switzer,A., McNeill,P.D. and Clapper,J.D.  
Compositions and methods for the therapy and diagnosis of lung  
Cancer  
Patent: WO 0204514-A 981 17-JAN-2002;  
CORIXA CORPORATION (US)

JOURNAL

TITLE

AUTHORS

FEATURES  
source  
1. .550  
Location/Qualifiers  
/Organism="Homo sapiens"  
/db\_xref="taxon:9606"

BASE COUNT 155 a 113 c 121 g 161 t

ORIGIN

Query Match 100.0%; Score 20; DB 6; Length 550;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTTCAACGAGATGTCATTC 20  
195 TTTCAACGAGATGTCATTC 214

Db 195 TTTCAACGAGATGTCATTC 214

RESULT 4  
HUMPAK/c 3052 bp mRNA linear PRI 10-JUN-2002

LOCUS HUMPAK/c

DEFINITION Homo sapiens focal adhesion kinase mRNA, complete cds.

ACCESSION L05186

VERSION L05186.1 GI:182394

KEYWORDS focal adhesion kinase; protein-tyrosine kinase.

SOURCE

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 3052)  
Andre,E. and Becker-Andre,M.  
Expression of an N-terminally truncated form of human focal  
adhesion kinase in brain  
Biochem. Biophys. Res. Commun. 190 (1), 140-147 (1993)

JOURNAL

MEDLINE 93135758

PUBMED 8422239

FEATURES  
source  
1. .3052  
Location/Qualifiers  
/Organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/tissue\_type="brain"  
/dev\_stage="fetal"  
15. .2654  
/EC\_number="2.7.1.1"  
/function="protein/tyrosine kinase"

CDS





RAB; annotation: HRI and RAB.  
FEATURES  
source  
1..3236  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="BRHIP2012923"  
/tissue\_type="hippocampus"  
/clone\_lib="BRHIP2"  
/note="cloning vector: pME18SFL3"  
13..2175  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="BAC04470.1"  
/db\_xref="GI:21754176"  
/translation="MADLIDGICRLVNGTSQSFITRPQKEGRALPSIPKLANSEKQG  
MRTAVSVSEDDVAEIIIDEDYTMPKSYGIDARDYEIQREKIELGRIGESQFG  
DVHOGIYMSPENPALAVAIKTCNKCTSDVRKFGQELTYRQPHPHIVKLGIVTE  
NPWIIIMELCTGELRSFLQVRKYSILDASLLVAYQSTALAYLESKRFVRIIDAR  
NVLVSNDVCVLGDFGLSRVMDSTYYKASKKLPIKMAPEISINFRFSTASDVMMF  
GVCMMELIMHGYPQGVKNNVIRIENGERELPMPNCPPTLYSLMTKCAVDPSSR  
PRTFLKQKQSTILLEEKAKQDEERKMRERQATVSMDSGSDAPPRKSPGTPSPR  
SSEGFYPSPPQHWQTNHYQVSGYPSHGIYMAAGSIYPGQASLLDQTSWMHPEIA  
MPOPNVEDSTVDLRGIGQVLPFLHMEERLIROQEMEDQRMLEKERFLKPDVRLS  
RGSIDREDSLOGPIGNQHIYQVGPKPAPPKPPRGAAGHLGSLASISSPADSY  
NEGVKLOPQOETSPPPTANLDRENDKYENVTGLVAVIEMSKIQAPPEEYVPMVKE  
VGLARTILATYDERTPLIPASTHEIEMAKQLNSDGLINKMKLAQOYVMTSLQ  
EYKQMLTPAHLAVDAKVLVDVIOAKRLKMLGQTRPH"  
BASE COUNT 961 a 697 c 741 g 837 t  
ORIGIN  
Query Match 100.0%; Score 20; DB 9; Length 3236;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TTTCAACGAGATGTCATTG 20  
|||||  
DB 2313 TTTCAACGAGATGTCATTG 2294  
RESULT 7  
LOCUS AR116314 3791 bp DNA linear PAT 16-MAY-2001  
DEFINITION Sequence 1 from patent US 6133031.  
ACCESSION AR116314  
VERSION AR116314.1 GI:14096636  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE  
1 (bases 1 to 3791)  
Monta, B.P. and Gaarde, W.A.  
TITLE Antisense inhibition of focal adhesion kinase expression  
JOURNAL Patent: US 6133031-A 1 17-OCT-2000;  
FEATURES  
source 1..3791  
/organism="unknown"  
BASE COUNT 1110 a 849 c 924 g 908 t  
ORIGIN  
Query Match 100.0%; Score 20; DB 6; Length 3791;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TTTCAACGAGATGTCATTG 20  
|||||  
DB 3529 TTTCAACGAGATGTCATTG 3510  
RESULT 8  
LOCUS HUMPAKX/c 3791 bp mRNA linear PRI 31-DEC-1994  
DEFINITION Human focal adhesion kinase (FAK) mRNA, complete cds.  
ACCESSION L13616

VERSION L13616.1 GI:439874  
KEYWORDS focal adhesion kinase; tyrosine kinase.  
SOURCE Homo sapiens (tissue library: lambda fcl1) T-lymphocyte cDNA to mRNA.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 3791)  
Whitney, G.S., Chan, P.Y., Blake, J., Cosand, W.L., Neubauer, M.G.,  
Aruffo, A. and Kanner, S.B.  
AUTHORS Human T and B lymphocytes express a structurally conserved focal  
adhesion kinase, pp125FAK  
JOURNAL DNA Cell Biol. 12 (9), 823-830 (1993)  
MEDLINE 94030617  
PubMed 7692878  
COMMENT 1478..2260 tyrosine kinase domain.  
FEATURES  
source  
1..3791  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/tissue\_type="T-lymphocyte"  
/tissue\_lib="lambda fcl1"  
1..3791  
/gene="FAK"  
1..232  
/gene="FAK"  
/note="putative"  
233..3391  
/gene="FAK"  
/note="putative"  
/codon\_start=1  
/product="focal adhesion kinase"  
/protein\_id="AAA58469.1"  
/db\_xref="GI:439875"  
/translation="MAAAVYLDPMNLNPNSSKTHLGTGERSPGAMERYLVKPHYE  
SNSERTWASIIIRHGADTVRGIIQKIVDSHKVYVACYGFRLSHLSEBVMILAVM  
GVSREKVELAPPEEMKVELRIYRYPKGLNPFQEDKDLNPFYQVSDVWLETA  
DQNDQELALKGCEIETRSYEMMGNALEKSYEVLEKVGAKRPFKSLDSVYKAK  
TLKRLIQOTFRQFANLNRESILKFFETLSVTRFDEKCRKALGSSWIIISVELAIGP  
EEGISYLDKGCNPTHLADFTQVQTIQVNSSEKDRKGMQLKIAAGPEPLVYAPSL  
TIAENMADLIDGICRLVNGTSQSFIIIRPQKEGRALPSIPKLANSEKQMRTHAVSV  
ETDDYAEIIIDEDYTMPSTROYEIORERIELGRCIQEGQFGDVHOGIYMSPENPALA  
VAITCKNCTSDSREKELOEALTMRFQDHPHIVKLGIVTEPNWIIIMELCTGELR  
SFLOVRKYSJDSLILYAYQLSTALAYLESKRFVRIIDARVLSNDVCYKLGDFG  
LSRYMEDSTYYKASKGKLPIKMAPEISINFRFTSADVMFVCMMELIMHGKPFQ  
GKXKNDVIGRIENGERLPMPCNPPTLYSLMTKCAVDPSSRPFELKQKQSTILLEE  
EKAQDEERKMRERQATVSMDSGSDAPPRKSPGTPSPRPFELKQKQSTILLEE  
HYVSGYPGSHGIIYMAAGSIYPGQASLLDQTSWMHPEIRIAMQPNVEDSTVDLRG  
IGQVLPFLHMEERLIROQEMEDQRMLEKERFLKPDVRLSRSIDREDSLOGPIG  
NQHIVQPVKQDPAPAPPKPPRGAAGHLGSLASISSPADSYNEGVKLOPQOETSPPPT  
ANDRSNDKYENVTGLVAVIEMSKIQAPPEEYVPMVKEVGLARTILATYDERTPLIP  
ASTHREIEMAKQLNSDGLINKMKLAQOYVMTSLQOEVKQMLTPAHLAVDAKVLVDV  
IOAKRLKMLGQTRPH"  
3' UTR  
3390..3791  
/gene="FAK"  
BASE COUNT 1110 a 849 c 924 g 908 t  
ORIGIN  
Query Match 100.0%; Score 20; DB 9; Length 3791;  
Best Local Similarity 100.0%; Pred. No. 22;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TTTCAACGAGATGTCATTG 20  
|||||  
DB 3529 TTTCAACGAGATGTCATTG 3510  
RESULT 9  
LOCUS HSM804272/c 4431 bp mRNA linear PRI 10-JUL-2002  
DEFINITION Homo sapiens mRNA; cDNA DKFZp66600110 (from clone DKFZp66600110).  
ACCESSION AL832961  
VERSION AL832961.1 GI:21733549

```

KEYWORDS
SOURCE
ORGANISM
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 4431)
AUTHORS
Ansorge,W., Wirkner,U., Mewes,H.W., Weil,B. and Wiemann,S.
TITLE
Direct Submission
JOURNAL
Submitted (09-JUL-2002) 1, D-85764 Neuberberg, GERMANY
COMMENT
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@kfz-heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFZp66600110) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.
FEATURES
Location/Qualifiers
source
1..4431
/organism="Homo sapiens"
/db xref="taxon:9606"
/clone="DKFZp66600110"
/tissue_type="stomach"
/clone_lib="666 (synonym: hsto2). Vector pSport1; host
DH10B; sites NotI + SalI"
/dev stage="adult"
polya signal
4383..4388
polya_site
4399
BASE_COUNT 1325 a 959 c 1023 g 1124 t
ORIGIN
Query Match 100.0%; Score 20; DB 9; Length 4431;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTCAACACGATGTCATTC 20
|||||
Db 3479 TTCAACACGATGTCATTC 3460
|||||

RESULT 10
AX281683/c
LOCUS
DEFINITION
Sequence 92 from Patent WO0177389.
ACCESSION
AX281683
VERSION
AX281683.1 GI:16608934
KEYWORDS
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS
Shiffman,D., Somogyi,R., Lawn,R., Seilhamer,J.J., Porter,G.J.,
Mikita,T. and Tai,J.
TITLE
Genes expressed in foam cell differentiation
JOURNAL
Patent: WO 0177389-A 92 18-OCT-2001;
Incyte Genomics, Inc. (US)
FEATURES
Location/Qualifiers
source
1..4481
/organism="Homo sapiens"
/db xref="taxon:9606"
/notes="Incyte ID No: 242114.16"
BASE_COUNT 1326 a 974 c 1045 g 1136 t
ORIGIN
Query Match 100.0%; Score 20; DB 6; Length 4481;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTCAACACGATGTCATTC 20
|||||

Db 3558 TTCAACACGATGTCATTC 3539
|||||

RESULT 11
AC067931
LOCUS
DEFINITION
Homo sapiens chromosome 8 clone RP11-642A1 map 8, *** SEQUENCING IN
PROGRESS ***, 1 ordered piece.
ACCESSION
AC067931
VERSION
AC067931.15 GI:22122968
KEYWORDS
HTG: HTGS_PHASE2; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 206467)
AUTHORS
Birren,B., Nusbaum,C. and Lander,E.
TITLE
Homo sapiens chromosome 8, clone RP11-642A1
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 206467)
AUTHORS
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
Campiano,A., Castle,A., Choquel,Y., Colangelo,M., Collins,S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczyk,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrum,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neil,D., Olivari,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tessfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (27-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 206467)
AUTHORS
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choquel,Y., Collymore,A.,
Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrum,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schupback,K., Seaman,S., Rogov,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Tessfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (06-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 6, 2002 this sequence version replaced gi:21450497.
All repeats were identified using RepeatMasker:
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research

```

Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
----- Project Information  
Center project name: L10082  
Center clone name: 642\_A\_1  
-----

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 1 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* provided by the submittor.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
1 206467: contig of 206467 bp in length.

FEATURES  
source  
1. 206467  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/map="8"  
/clone="RP11-642A1"  
/clone\_11b="RPC1-11 Human Male BAC"  
BASE COUNT 57619 a 50155 c 48369 g 50320 t 4 others  
ORIGIN

Query Match 100.0%; Score 20; DB 2; Length 206467;  
Best Local Similarity 100.0%; Pred. No. 19;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTTCACCGAGTGTGTCATTC 20  
|||||  
Db 135959 TTTCACCGAGTGTGTCATTC 135978

RESULT 12  
RMEI32004 3974 bp DNA linear BCT 25-MAR-1999  
LOCUS Rhizobium melioli c1p and leuA genes, ORF83 and ORF142.  
DEFINITION  
AJ132004  
ACCESSION  
VERSION AJ132004.1 GI:4539548  
KEYWORDS 2-Isopropylmalate synthase; c1p protease; c1p gene; leuA gene;  
leucine biosynthesis; ORF142; ORF83; proteolytic subunit.  
SOURCE Sinorhizobium melioli.  
ORGANISM Sinorhizobium melioli.  
Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
Rhizobiaceae; Sinorhizobium.  
1 (bases 1 to 3974)  
Sanjuan-Pinilla,V.M., Olivares,J. and Sanjuan,J.  
The Rhizobium melioli leuA gene is essential for symbiosis  
Unpublished  
2 (bases 1 to 3974)  
Sanjuan,J.  
Direct Submission  
Submitted (24-FEB-1999) Sanjuan J., Microbiologia del Suelo y  
Sistemas Simbioticos, Estacion Experimental del Zaidin-CSIC,  
Profesor Albarada 1, E-18008 Granada, SPAIN  
location/Qualifiers  
1.3974  
/organism="Sinorhizobium melioli"  
/strain="GR4"  
/db\_xref="taxon:382"  
136..723  
/gene="c1p"  
136..723  
/gene="c1p"  
/codon\_start=1  
/transl\_table=1  
/product="C1p protease proteolytic subunit"  
/protein\_id="CAB39976.1"  
/db\_xref="GI:4539549"  
/translation="MRNDQDEKTEPLPGKETANLFKRSIFITGTTQELAQXV

gene  
CDS

CSOLVALAASDDDIRLVNSPGGVESGDSIHDMIKFVKPVKTITGNGVASAGALI  
VYAPKSORCLPNTREILHQSGCTGMSDIEIQAREIILKNMERLNRIFSEATGCP  
VDKAKQIDRDYMLGABEAKYGLVSNITVTSIABI"  
1194..2903  
/gene="leuA"  
1194..2903  
/gene="leuA"  
/function="leucine biosynthesis"  
/codon\_start=1  
/transl\_table=1  
/product="2-Isopropylmalate synthase"  
/protein\_id="CAB39977.1"  
/db\_xref="GI:4539550"

CDS

/translation="MILKSHSIXKGMPEAAVKQPYPOIVLPDRTMPSKATTEPNC  
SVDRDGNQALVDPMGHDKRKARNPHILLMDGFKETIEGFPASQDTYDFARMCYEGN  
VSEBVSQVLAQCPBELIATFEALBEAHNPYVHFVSTSELQRRVVGKQVAKIKOI  
ADTAAKMITDMAAKAGGGRFEYSPESEFTLEALAEICNAVEIYRPNADNGLIIN  
LPSTVEMATPNIYADQIEMMCRLNDRENIIVSIAPHNDGSGTIAATDGLMAGADRV  
EGTLFNGERTGNVDVVTLLNMFQGVDPKLDGSDIERIKVEYENQMWYIPRRHY  
VGEIVYTAFGSGSHDAINKGMKAKQANKPTWEYPIPIPRDVGSRYEAMIRINSOS  
GKGGIATYLGEDVGINLPRLNQLEFREVEPRITDEGKELPSKIHQRFIESVYQPG  
ARIKFVDDHTYYPAGEHKGRLVVAAEITDGETRQIEGKGPIDGFTNASTIYGIEL  
SVADYSEHSIQHGSMAAIAIYVEVEYPGKLFVGIINTNIVAASTLEAIVSAANRVLIV  
VGK"  
3041..3292  
/note="ORF83"  
/codon\_start=1  
/transl\_table=1  
/product="hypothetical protein"  
/protein\_id="CAB39978.1"  
/db\_xref="GI:4539551"  
/translation="MALYIKDPTVDRMAEKLQERTGRTKDANRIALQHELDVDEBE  
IPREKJLAALRQKANDRLGPPVHGIDMKLDELMEGE"  
3289..3717  
/note="ORF142"  
/codon\_start=1  
/transl\_table=1  
/product="hypothetical protein"  
/protein\_id="CAB39979.1"  
/db\_xref="GI:4539552"  
/translation="MSEFVDAVIAIINEBGFEELEKRLSDADGKLCVSPLYRFEAV  
AALPRLRLITKGRADRSIDLGEARELVSDFIQLSASEVTIDSHSTGVRALDAMARYG  
KVAGHPALNLDGCFAYAAAKESGLTIYKGNDSQDLDG"

CDS

BASE COUNT 828 a 1198 c 1205 g 743 t

Query Match 92.0%; Score 18.4; DB 1; Length 3974;  
Best Local Similarity 95.0%; Pred. No. 1.5e+02;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTTCACCGAGTGTGTCATTC 20  
|||||  
Db 2159 TTTCACCGAGTGTGTCATTC 2178

RESULT 13  
SMES591790/c 323450 bp DNA linear BCT 05-JUL-2002  
LOCUS Sinorhizobium melioli 1021 complete chromosome, segment 9/12.  
DEFINITION  
AL591790 AL591688  
ACCESSION  
VERSION AL591790.1 GI:15075230  
KEYWORDS  
SOURCE Sinorhizobium melioli.  
ORGANISM Sinorhizobium melioli.  
Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
Rhizobiaceae; Sinorhizobium.  
1 (bases 1 to 323450)  
Capela,D., Barloy-Hubler,F., Gouzy,J., Bothe,G., Ampe,F., Batur,J.,  
Boistard,P., Becker,A., Bouty,M., Cadieu,B., Dreano,S., Gloux,S.,  
Godrie,T., Goffeau,A., Kahn,D., Kiss,E., Lelaure,V., Maesuy,D.,  
Pohl,T., Portetalle,D., Puhler,A., Purnelle,B., Ramsperger,U.,  
Renard,C., Thebaud,P., Vandenbol,M., Weidner,S. and Galibert,F.  
Analysis of the chromosome sequence of the legume symbiont

JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE JOURNAL	Sinorhizobium meliloti strain 1021 Proc. Natl. Acad. Sci. U.S.A. 98 (17), 9877-9882 (2001) 21396507 11481430 2 (bases 1 to 323450) Gouzy, J. Direct Submission Submitted (26-JUL-2001) Gouzy J., Submitted on behalf of the MELILO EU Consortium	gene CDS	KAHLSGAEGETCALLDRRAGTIQPLAYARGIARAALAAGAEIFTDTPLLAASROGDL WNLKTPTGTVTARHVLATNAYGSLTVGPWKYEOELTILPYFOFATNPLPDNVAAR ILPERGAWDITGLVMTSPMRDRNRLIFVSGICRLDAIAAGTHRAFAARSVKLFPPYIG DPRFEHWDGRIGMTTNNLPAHVLAPNVVSGYSGVGIAPGTVFGRALARHVTGDT SAIPLAETPVTDPWRTTKSAFYHAGAQAQKHFIDKRF" complement (3024..3740) /gene="SMC01579" complement (3024..3740) /gene="SMC01579" /function="miscellaneous; hypothetical/global homology" /note="Product confidence : hypothetical Gene name confidence : hypothetical predicted by Codon_usage predicted by Homology predicted by FrameD" /codon_start=1 /evidence=not experimental /transl_table=11 /product="CONSERVED HYPOTHETICAL PROTEIN" /protein_id="CAC46789.1" /db_xref="GI:15075233" /db_xref="SPTREMBL:Q92NJ2" /translation="MAVITFANAKGGAGTKTAALILSTELARQGNVVVLDADPORWI TSMSEVSRVANLEVISHTVPASLPHCHIRELGEADFVIDLAGAKDAIVALALGSLD HVLIPVQCANDARGAVQILLEIRHIGEKARVRINHVSVLTRVNSLVTTRALQTKAL LASRGVSLDTPIVERVAYREIFECGGTLQMDPNRVSNLKDARENAYALAAAEVQNL PVTARRALMSRLSALPRAA" complement (3949..5094) /gene="SMC01582" complement (3949..5094) /gene="SMC01582" /EC_number="1.1.1.-" /function="small molecule metabolism" /note="Product confidence : putative Gene name confidence : hypothetical predicted by Codon_usage predicted by Homology predicted by FrameD" /codon_start=1 /evidence=not experimental /transl_table=11 /product="PUTATIVE ALCOHOL DEHYDROGENASE PROTEIN" /protein_id="CAC46790.1" /db_xref="GI:15075234" /db_xref="SPTREMBL:Q92NJI" /translation="MTITANWSYPTAVKFGAGRIKELADHCKALGIKKPLVTVDRGLA PMAITQQALDILEAGLGRAIFADVDPNDNLEAGVKASRDGHDGCVAFGGSGGL DLKCVAFMAQGTFRVDFEIGDWWTRASVEGIIAPIVAVPTTAGTSGEVGRASVITN SASHVKVFIHFPIKFLPGVTICDPELTVMGPVKVITAGTQMDAFAPHCLEAYSGPFYHPMS AGIALLEGMLVKYELPRAYKDGADLEARNMMAAMGAVAFQKGLGAIHSLSHFVGA IYNTHGNTNAVMPVPLNRSALIEEKIGRAAYLGIAGGDFDGYDYLRLREELGV PKLSALGVGTDRIDEMAEMAIVDPTAGNPNVELTDLAAEKLFAECI" complement (5091..6476) /gene="SMC01588" complement (5091..6476) /gene="SMC01588" /EC_number="1.2.1.-" /function="small molecule metabolism" /note="Product confidence : putative Gene name confidence : hypothetical predicted by Codon_usage predicted by Homology predicted by FrameD" /codon_start=1 /evidence=not experimental /transl_table=11 /product="PUTATIVE ALDEHYDE DEHYDROGENASE PROTEIN" /protein_id="CAC46791.1" /db_xref="GI:15075235" /db_xref="SPTREMBL:Q92NJO" /translation="WTMIRCVSDGEVYASRPAIPLEMAQVAHARLAOKAWARRP LDERVKVLVAGVARNLNMVDVVPPELAGMGRPVRYGGEFGFNERSYVASIRADAL KPLVVEESDRFERRIAREPHGVFVFIAPWNPYMTAINTVAPMAGNTVILKHASQT
JOURNAL MEDLINE PUBMED REFERENCE AUTHORS TITLE JOURNAL	MELILO EU Consortium: Laboratoire de Biologie Molculaire des Relations Plantes-Microorganismes, UMR215-CNRS-INRA, BP27, F-31326 Castanet, France, Laboratoire de Genetique et Developpement UMR6061-CNRS, France, Laboratoire de Genetique et Developpement UMR6061-CNRS, Faculte de Medecine, 2 avenue du Pr. Leon Bernard, F-35043 Rennes, France, GATC GmbH, Fritz-Arnold-str. 23, D-78467 Konstanz, Germany, Universitaet Bielefeld, Biologie IV (Genetik) Universitaetstr 25, D-33615 Bielefeld, Germany, Unite de Biochimie physiologique, Universite Catholique de Louvain, Place Croix du Sud 2, Bte 20, B-1348 Louvain-la-Neuve, Belgium, Unite de Microbiologie, Faculte des Sciences Agronomiques de Gembloux, Avenue Marechal Juin 6, B-5030 Gembloux, Belgium. E-mail:Jerome.Gouzy@coulouse.inra.fr http://sequence.toulouse.inra.fr/meliloti.html. Location/Qualifiers 1..323450 /organism="Sinorhizobium meliloti" /strain="1021" /db_xref="taxon:382" 155..1519 /gene="SMC01575" 155..1519 /gene="SMC01575" /function="miscellaneous; hypothetical/partial homology" /note="Product confidence : hypothetical Gene name confidence : hypothetical predicted by Codon_usage predicted by Homology predicted by FrameD" /codon_start=1 /evidence=not experimental /transl_table=11 /product="HYPOTHETICAL PROTEIN" /protein_id="CAC46787.1" /db_xref="GI:15075231" /db_xref="SPTREMBL:Q92K34" /translation="MTPATRTGLSLFAALALMAMELAPANAQDAYGYWGCGDVMLVT PGEIDILYIPGEAEHAMDRGRITVLDPWGNIVATVVPNDYGRGRRYGRDGY PPDRDGYGSPGEFTGAIPEVRDIAPAPEREDLPNSLPSLSDBREAAVPOYDDP LAQMPAPMTVTKGSAEIAALQVLDREGSPGVIDGKWSNVTKAIEAWQAQGTET LDPNNTEDILRLRNFNGGIPITYITITADAAGPFVVASIPEDTAHKAQPLHLSFTSVT EMLGEKPHDEAYLRELNPGVDPSIPGTTIKVNPNGPKKGVARIADKARKQVLA DEAGKLIAPSTIGSSDTPSPSGTVHVERIAFDPGYTYNPKINFIQQGANDRILOLQ GNGPVGTVWIALSKPTYGIHGTPSPKIGTKQSHGCVRLTNWDATLGLKMYSTGTVV EFVD" 1707..2996 /gene="SMC01576" 1707..2996 /gene="SMC01576" /function="small molecule metabolism" /note="Product confidence : putative Gene name confidence : hypothetical predicted by Codon_usage predicted by Homology predicted by FrameD" /codon_start=1 /evidence=not experimental /transl_table=11 /product="PUTATIVE OXIDOREDUCTASE PROTEIN" /protein_id="CAC46788.1" /db_xref="GI:15075232" /db_xref="SPTREMBL:Q92NJ3" /translation="WATERALPNLWHTAPAPRTAPLAGDLTVEVAIVGGGFTGLSA ALHLAETGIRTAVERIARMIGFGSGNRNVLNAGMWVQPDILTATLGAGNRLDLDEL GDGPFVYDLVAKHGIECEAVRNGTLHMSVGAEGLKEITRREARQKKGAPVEVLSAE	gene CDS	gene CDS

```

ILGEMRVAFIEBAGVPADVQONFLDHTTAALIAKSPFINTSGVEGGRSTERA
AAGFTTGLEIGCDPGYVMEADADLDAVDTLMDGATYNSGQCCGIERIYVESLY
DAPEKSVAAVNSYKGNPLDPEITLGMANKRFAATVRNOVADLISGAALIDPKL
FPDDGAYVLAPOVLVDVDMKEMFERETFGAVIMKYSKDAEALIELMNSRVLDTA
SLMTDARARARISGRDLETGVFMNRADYLDPALCMTGVKEIRGSGSLVGFNLTR
EKSHYLUKVYA"
gene
complement (6674, .8038)
/gene="SMC01594"
complement (6674, .8038)
/gene="SMC01594"
/function="miscellaneous; hypothetical/global homology"
/notice="product confidence: hypothetical
gene name confidence: hypothetical
predicted by Codon usage
predicted by Homology
predicted by Framed"
/codon_start=1
/evidence=not experimental
/translation="transl_table=11
/product="CONSERVED HYPOTHETICAL PROTEIN"
/protein_id="CAC46792.1"
/db_xref="GI:15075236"
/db_xref="SPTREMBL:O92NI9"
/translation="MSYFBEELKEKEDVAAGRIDTVLACQVDMQGRUMGRFAHYVES
AMKETHSCNYLATDMEETVPGVATISWEKQYDITKPDISTLRRIPIHEGTLVL
CDMLDHDHAEVPHSPRAILKKQVRLAMGKAWASELEFDFQSYDARISGR
DLQALAGYNEDEYHIFQTKREBDMRAIRNGQAGI PVENSKGESAQOEINRYAD
LUTMADRHAITIKNGCKEITAMWGKAITFLAKWNSAGSSSHISQSKDGEYPLF
DKNOYGSSELRRHYVAGOLAHASEVTFLAPYINSYKRPMAQGTAPAPKWSKONT
AGYRLCGSGSKAIRIECRVGSGLDNPYLAFLALIAAGIAGIENKRELEAPVGDAYOG
KEVERIPHTLRHAGALSQSKMLRAFAEBEVVDHYHAHEWQOEYDRRTDWEVARG
FERA"
gene
complement (8131, .9678)
/gene="SMC01597"
complement (8131, .9678)
/gene="SMC01597"
/function="cell processes; transport of small molecules;
amino acids, amines, peptides"
Query Match 92.0%; Score 18.4; DB 1; Length 323450;
Best Local Similarity 95.0%; Pred No. 1.3e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 TTTCACCCAGATGTCATTC 20
Db 183366 TTCCAACGAGATGTCATTC 183347
RESULT 14
AC102801 119671 bp DNA linear PRI 01-MAY-2002
LOCUS Homo sapiens chromosome 8, clone RP11-960P16, complete sequence.
DEFINITION AC102801
AC102801
AC102801.2 GI:20377028
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 119671)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Bouknight,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dode,S., Fato,S.,
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Grinde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heard,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,

```

```

Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Margus,N., Mathews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheters,R., Meldrim,J.,
Menueh,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunhphang,P., Pierre,N., Pollard,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R.,
Seaman,S., Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Teefaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigglio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02144, USA
3 (bases 1 to 119671)
REFERENCE
JOURNAL TITLE
AUTHORS
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
Bouknight,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dode,S.,
Fato,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Grinde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Margus,N., Mathews,C.,
McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Menueh,L.,
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicoll,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunhphang,P., Pierre,N., Pollard,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R.,
Seaman,S., Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Teefaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigglio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (01-MAY-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02144, USA
On May 1, 2002 this sequence version replaced gi:117060911.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WITR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L19703
Center clone name: 960_P_16
-----
SP6 end overlaps project AC011632 (L2781) by 63.5 kilobases; T7 end
overlaps project AC100782 (L19699) by 47.5 kilobases. We are
submitting only the middle 119.7 kilobases of the project.
FEATURES
. source
Location/Qualifiers
1. 119671
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8"
/clone="RP11-960P16"
/clone_id="RP11-11 Human Male BAC"
1. 115
/rpt_family="AluY"
repeat_region 295..366
repeat_region 403..447
rpt_family="AT_rich"

```

```

repeat_region complement(537..897)
repeat_region /rpt_family="MLT1A1"
repeat_region complement(954..1055)
repeat_region /rpt_family="L1MC2"
repeat_region complement(1059..1320)
repeat_region /rpt_family="AluJb"
repeat_region complement(1321..1816)
repeat_region /rpt_family="L1MC2"
repeat_region 2593..2597
repeat_region /rpt_family="L2"
repeat_region 2938..2991
repeat_region /rpt_family="(TCCCC)n"
repeat_region complement(3143..3447)
repeat_region /rpt_family="L1MC4a"
repeat_region complement(4363..4605)
repeat_region /rpt_family="AluJb"
repeat_region complement(4609..4939)
repeat_region /rpt_family="L1MC4a"
repeat_region 4957..5281
repeat_region /rpt_family="AluSx"
repeat_region 5632..5895
repeat_region /rpt_family="L1PA13"
repeat_region complement(6877..7093)
repeat_region /rpt_family="MER4D"
repeat_region 7094..7136
repeat_region /rpt_family="(TC)n"
repeat_region complement(7137..7156)
repeat_region /rpt_family="MER4D"
repeat_region 7157..7429
repeat_region /rpt_family="AluSg"
repeat_region complement(7430..7459)
repeat_region /rpt_family="MER4D"
repeat_region complement(7475..7576)
repeat_region /rpt_family="MER4D"
repeat_region 7580..7731
repeat_region /rpt_family="L2"
repeat_region complement(8168..8342)
repeat_region /rpt_family="HAL1"
repeat_region 8648..8676
repeat_region /rpt_family="(TTCTCC)n"
repeat_region complement(8760..9400)
repeat_region /rpt_family="HAL1"
repeat_region 9418..9451
repeat_region /rpt_family="AT_rich"
repeat_region complement(9494..9610)
repeat_region /rpt_family="FLAM_A"
repeat_region complement(9662..9939)
repeat_region /rpt_family="HAL1"
repeat_region complement(10387..10511)
repeat_region /rpt_family="MER91C"
repeat_region 10824..10881
repeat_region /rpt_family="(ATG)n"
repeat_region 11814..12716
repeat_region /rpt_family="MER83B-int"
repeat_region 12742..13734
repeat_region /rpt_family="MER83B-int"
repeat_region 13731..14158
repeat_region /rpt_family="MER83B-int"
repeat_region 14159..14189
repeat_region /rpt_family="(CAA)n"
repeat_region 14190..16843
repeat_region /rpt_family="MER83B-int"
repeat_region 16846..17345
repeat_region /rpt_family="L1P"
repeat_region complement(17347..17514)
repeat_region /rpt_family="AluSx"
repeat_region 18137..18161
repeat_region /rpt_family="AT_rich"
repeat_region 18223..18249
repeat_region /rpt_family="(T)n"
repeat_region 19179..19213
repeat_region /rpt_family="AT_rich"
repeat_region 19357..19380
/rpt_family="(TTTG)n"
complement(19719..19861)
/rpt_family="L2"
complement(19894..20075)
/rpt_family="MER69A"
20383..20739
/rpt_family="MLT1A1"
complement(21160..21300)
/rpt_family="MIR"
complement(21406..21469)
/rpt_family="L1PA13"
complement(21478..21704)
/rpt_family="L2"
22711..23230
/rpt_family="MER1A"
complement(23983..24533)
/rpt_family="MLT1E"
24756..25122
/rpt_family="MSTA"
complement(26267..26376)
/rpt_family="MIR3"
complement(27660..27918)
/rpt_family="L2"
28096..28142
/rpt_family="AT_rich"
28466..28628
/rpt_family="Charlie1"
28633..28759
/rpt_family="Charlie1"
complement(29057..29178)
/rpt_family="MIR"

Query Match 90.0%; Score 18; DB 9; Length 119671;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TTCAACCAGATGTCATT 19
|||||
Db 35784 TTCAACCAGATGTCATT 35801

RESULT 15
AC084336/c
LOCUS
DEFINITION Homo sapiens chromosome 8, clone RP11-625F15, complete sequence.
ACCESSION AC084336
VERSION AC084336.6 GI:17488656
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 156235)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 8, clone RP11-625F15
JOURNAL Unpublished
REFERENCE
2 (bases 1 to 156235)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bada,P., Boguslavskiy,L.,
Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
DeArrellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,
FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,
Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L.,
Illiev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRocque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G.,
MacDonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K.,
McPheters,R., Meldrim,J., Meneus,L., Mihova,T., Mlenga,V.,
Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,
O'Donnell,P., O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K.,
Pierre,N., Pisani,C., Pollara,V., Raymond,C., Rieback,M., Riley,R.,
Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,
Sougnuez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,

```

Straus, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,  
Tittell, A., Travers, M., Triggillo, J., Vassiliev, H., Viel, R., Vo, A.,  
Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J.,  
Zimmer, A. and Zody, M.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

Submitted (25-OCT-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 156235)  
Birtten, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N.,  
Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhvalter, B.,  
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,  
Choepe, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,  
Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S.,  
Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gaddys, S.,  
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,  
Hagos, B., Hearford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,  
Jones, C., Kamat, A., Karatas, A., Kells, C., Lacroque, K.,  
Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,  
Maclean, C., Macdonald, P., Major, J., Margis, N., Matthews, C.,  
McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrum, J.,  
Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C.,  
Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,  
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,  
Raymond, C., Reta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,  
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R.,  
Seaman, S., Severy, P., Sougnuez, C., Spencer, B., Stange-Thomann, N.,  
Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Testaye, S.,  
Theodore, J., Topham, K., Travers, M., Travis, N., Triggillo, J.,  
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,  
Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

Submitted (31-JUL-2001) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
4 (bases 1 to 156235)  
Birtten, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N.,  
Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhvalter, B.,  
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,  
Choepe, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,  
Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S.,  
Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gaddys, S.,  
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,  
Hagos, B., Hearford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,  
Jones, C., Kamat, A., Karatas, A., Kells, C., Lacroque, K.,  
Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,  
Maclean, C., Macdonald, P., Major, J., Margis, N., Matthews, C.,  
McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrum, J.,  
Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C.,  
Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,  
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,  
Raymond, C., Reta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,  
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R.,  
Seaman, S., Severy, P., Sougnuez, C., Spencer, B., Stange-Thomann, N.,  
Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Testaye, S.,  
Theodore, J., Topham, K., Travers, M., Travis, N., Triggillo, J.,  
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,  
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE  
JOURNAL  
COMMENT

Submitted (11-DEC-2001) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Dec 11, 2001 this sequence version replaced gi:15042917.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
Project Information  
Center project name: L10933  
Center clone name: 625\_F\_15  
Location/Qualifiers

source	1..156235
repeat_region	/organism="Homo sapiens"
repeat_region	/cd_xref="taxon:9606"
repeat_region	/chromosome="8"
repeat_region	/map="8"
repeat_region	/clone="RP11-625F15"
repeat_region	/clone_lib="RPCT-11 Human Male BAC"
repeat_region	complement(106..401)
repeat_region	/rpt_family="Alu5x"
repeat_region	complement(1440..1685)
repeat_region	/rpt_family="MIR"
repeat_region	5277..5305
repeat_region	/rpt_family="AT_rich"
repeat_region	5475..5549
repeat_region	/rpt_family="L2"
repeat_region	6414..6805
repeat_region	/rpt_family="MLT1L"
repeat_region	7240..7466
repeat_region	/rpt_family="MIR"
repeat_region	7468..7886
repeat_region	/rpt_family="LTR35"
repeat_region	8012..8287
repeat_region	/rpt_family="Alu5x"
repeat_region	8289..8313
repeat_region	/rpt_family="TTAA)n"
repeat_region	complement(8425..8607)
repeat_region	/rpt_family="L1ME"
repeat_region	complement(8630..8990)
repeat_region	/rpt_family="L1ME"
repeat_region	9173..9464
repeat_region	/rpt_family="L1PB3"
repeat_region	11763..11799
repeat_region	/rpt_family="AT_rich"
repeat_region	11986..12230
repeat_region	/rpt_family="MLT1H"
repeat_region	complement(12437..12523)
repeat_region	/rpt_family="L1MA4"
repeat_region	complement(12604..12912)
repeat_region	/rpt_family="Alu5x"
repeat_region	complement(13057..13515)
repeat_region	/rpt_family="L1PA13"
repeat_region	17016..17434
repeat_region	/rpt_family="L1M4"
repeat_region	complement(17940..18086)
repeat_region	/rpt_family="MIR"
repeat_region	18132..18338
repeat_region	/rpt_family="MIR"
repeat_region	complement(18371..18430)
repeat_region	/rpt_family="MLT2B3"
repeat_region	18431..18728
repeat_region	/rpt_family="Alu5x"
repeat_region	complement(18729..19208)
repeat_region	/rpt_family="MLT2B3"
repeat_region	19252..19310
repeat_region	/rpt_family="L2"
repeat_region	complement(19429..19553)
repeat_region	/rpt_family="MER69"
repeat_region	complement(19585..19952)
repeat_region	/rpt_family="MER69"
repeat_region	23781..23801
repeat_region	/rpt_family="AT_rich"
repeat_region	23928..24244
repeat_region	/rpt_family="AluYb8"
repeat_region	complement(25563..25882)
repeat_region	/rpt_family="AluYb8"
repeat_region	complement(25893..26166)
repeat_region	/rpt_family="AluY"
repeat_region	26170..26242
repeat_region	/rpt_family="MER20B"
repeat_region	complement(26338..26565)
repeat_region	/rpt_family="MER20B"
repeat_region	27102..27951
repeat_region	/rpt_family="L1MC3"

FEATURES



```

repeat_region 27965. .28374
                /rpt_family="L1MC4a"
repeat_region 28375. .28446
                /rpt_family="MLT1F1"
repeat_region 28447. .29044
                /rpt_family="L1MC4a"
repeat_region complement(30668. .30969)
repeat_region /rpt_family="AluY" .31254)
repeat_region complement(31214. .31254)
repeat_region /rpt_family="L2"
repeat_region 32008. .32179
                /rpt_family="L1M4"
repeat_region complement(32354. .32734)
repeat_region /rpt_family="MLT1J2"
repeat_region 33062. .33193
                /rpt_family="AluSq/x"
repeat_region 33815. .33836
                /rpt_family="AT rich"
repeat_region complement(34016. .34092)
repeat_region /rpt_family="L2"
repeat_region complement(34089. .34506)
repeat_region /rpt_family="L2"
repeat_region complement(34594. .34765)
repeat_region /rpt_family="L1P1"

```

Query Match 90.0%; Score 18; DB 9; Length 156235;  
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;  
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 2 TTCAACCATGGTCATT 19
   |||||
Db 65383 TTCAACCATGGTCATT 65366

```

Search completed: January 12, 2003, 10:38:36  
 Job time : 2817 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2003, 09:40:15 ; Search time 253 Seconds  
(without alignments)  
178.024 Million cell updates/sec

Title: US-09-757-100B-18  
Perfect score: 20  
Sequence: 1 ttctaccagatggtcattc 20

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_101002.\*  
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	22	Human focal adhesi
2	20	100.0	20	24	Human focal adhesi
3	20	100.0	550	24	CDNA encoding lung
C 4	20	100.0	3791	22	Human focal adhesi
C 5	20	100.0	3791	24	Human focal adhesi
C 6	20	100.0	4052	23	Human prostate exp
C 7	20	100.0	4481	24	Human DNA sequence
8	18.4	92.0	707	22	Human neuroblastom
9	17	85.0	1311	22	Coriolus versicolor

10	16.8	84.0	244	22	ABAS1579	Human breast cell
11	16.8	84.0	244	22	ABA35555	Probe #15021 for g
12	16.8	84.0	244	22	AAK3716	Human bone marrow
13	16.8	84.0	400	22	ABA46482	Human breast cell
14	16.8	84.0	400	22	ABA26688	Probe #5154 for ge
15	16.8	84.0	400	22	AAK30708	Human bone marrow
16	16.8	84.0	528	22	AAK37562	Human bone marrow
17	16.8	84.0	528	22	AAI43434	Probe #12120 used
18	16.8	84.0	1681	14	AAQ34687	Babesia equi subcl
19	16.8	84.0	1911	24	ABL91702	Human polynucleoti
C 20	16.8	84.0	4544	10	AAAN90355	CDNA encoding plat
C 21	16.8	84.0	4544	17	AAAT34552	Platelet-derived g
C 22	16.8	84.0	5427	13	AAQ27447	Type B human plate
C 23	16.8	84.0	5602	23	AAAS84940	CDNA encoding novel
C 24	16.8	84.0	5719	10	AAAN90388	CDNA encoding huma
C 25	16.8	84.0	6390	23	AAAS79665	DNA encoding novel
C 26	16.8	84.0	6390	23	AAAS84936	DNA encoding novel
C 27	16.8	84.0	11036	20	AAAX13401	Enterococcus faeca
28	16.8	84.0	32191	22	AAAS26704	Human genomic DNA
C 29	16.8	84.0	49561	22	AAK82012	Human immune/haema
30	16.4	82.0	4094	21	AAZ51205	Human hypoxia regu
31	15.8	79.0	175	21	AAAC14051	Human secreted pro
32	15.8	79.0	476	24	ABL83730	Human ovarian canc
33	15.8	79.0	490	23	AAAS87979	DNA encoding novel
34	15.8	79.0	575	24	ABK78062	Bacillus clausii g
35	15.8	79.0	579	23	AAAS68415	DNA encoding novel
C 36	15.8	79.0	709	18	AAAX30750	Streptococcus pneu
37	15.8	79.0	822	23	AAAS7983	DNA encoding novel
C 38	15.8	79.0	860	21	AAA05546	Streptococcus pneu
C 39	15.8	79.0	1020	23	AAAS55500	Streptococcus pneu
C 40	15.8	79.0	1020	23	AAAS55900	Streptococcus pneu
41	15.8	79.0	1377	24	ABA92851	Yersinia pseudotub
42	15.8	79.0	2286	23	AAAS92528	DNA encoding novel
C 43	15.8	79.0	3643	23	ABL03334	Drosophila melanog
44	15.8	79.0	5304	23	ABL20423	Drosophila melanog
45	15.8	79.0	5358	23	ABL04503	Drosophila melanog

ALIGNMENTS

RESULT 1  
AAC65550  
ID AAC65550 standard; DNA; 20 BP.

XX AC AAC65550;

XX DT 12-FEB-2001 (first entry)

XX DE Human focal adhesion kinase antisense sequence #16.

XX KW Human; focal adhesion kinase; FAK; signal transduction; cancer;

XX KW embryonic development disorder; angiogenic disorder; wound healing;

XX KW antisense; phosphorothioate; ss.

XX OS Homo sapiens.

XX PN US6133031-A.

XX PD 17-OCT-2000.

XX PF 19-AUG-1999; 99US-0377310.

XX PR 19-AUG-1999; 99US-0377310.

XX PA (ISIS-) ISIS PHARM INC.

XX PI Monia BP, Gaarde WA;

XX DR WPI; 2001-006141/01.

XX PT New antisense compounds for inhibiting focal adhesion kinase expression, especially useful for inhibiting retinal

PT neovascularization, or for diagnosing and treating e.g. colon cancer -  
 XX  
 PS Claim 3; Column 23; 30pp; English.  
 XX  
 CC The present invention describes a number of phosphorothioate antisense  
 CC sequences to the human focal adhesion kinase (FAK) protein. This protein  
 CC is involved in integrin-mediated signal transduction, and is implicated  
 CC in cancer, particularly colon, breast and oral tumours, embryonic  
 CC development disorders, angiogenic disorders and wound healing. The  
 CC antisense sequences, including the one shown here, can be used in the  
 CC treatment of all of these.  
 XX  
 SQ Sequence 20 BP; 5 A; 5 C; 3 G; 7 T; 0 other;  
 XX  
 Query Match 100.0%; Score 20; DB 22; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 0.78;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TTTCACCGAGATGTCATTC 20  
 Db 1 TTTCACCGAGATGTCATTC 20  
 RESULT 2  
 AAS15398  
 ID AAS15398 standard; DNA; 20 BP.  
 XX  
 AC AAS15398;  
 XX  
 DT 29-JAN-2002 (first entry)  
 XX  
 DE Human focal adhesion kinase (FAK) antisense oligonucleotide IS15421.  
 XX  
 KW Human; focal adhesion kinase; FAK; melanoma; tumour metastasis;  
 KW cancer; angiogenic disorder; retinal neovascularisation; cytosstatic;  
 KW ophthalmological; antisense; phosphorothioate; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT modified\_base 1..20  
 FT /tag= a  
 FT /mod\_base= OTHER  
 FT /note= "phosphorothioate internucleotide linkages,  
 FT optionally bases 1-5 and 16-20 are  
 FT 2'-methoxyethoxy (2'-MOE) bases, where the  
 FT 2'-MOE cytosines are also 5-methylcytosines"  
 XX  
 PN US2001034329-A1.  
 XX  
 PD 25-OCT-2001.  
 XX  
 PF 09-JAN-2001; 2001US-0757100.  
 XX  
 PR 19-AUG-1999; 99US-0377310.  
 PR 13-JUL-2000; 2000WO-US18999.  
 XX  
 PA (MONI/) MONIA B P.  
 PA (GAAR/) GAARDE W A.  
 PA (NERO/) NERO P S.  
 XX  
 PI Monia BP, Gaarde WA, Nero PS;  
 XX  
 DR WPI; 2002-010103/01.  
 XX  
 PT Antisense suppression of Focal Adhesion Kinase expression for the  
 PT treatment of cancers of the breast, colon, mouth or skin (especially a  
 PT melanoma), and angiogenic disorders e.g. retinal neovascularisation -  
 XX  
 PS Claim 3; Page 12; 19pp; English.  
 XX  
 CC The present invention relates to novel antisense compounds which can be  
 CC used for modulating the expression of human focal adhesion kinase (FAK).

CC The antisense compounds comprise antisense oligonucleotides  
 CC (8-30 nucleotides in length) targeted to the 5'-untranslated region,  
 CC translational termination region or 3' untranslated region of a nucleic  
 CC acid molecule encoding FAK. The antisense oligonucleotides can be used  
 CC to inhibit the expression of FAK mRNA. The antisense oligonucleotides  
 CC represent potential chemotherapeutic agents in the treatment of melanoma  
 CC and the prevention of tumour metastasis. The antisense compounds and  
 CC methods are useful for treating diseases associated with overexpression  
 CC or constitutive activation of FAK. Such diseases include cancers of the  
 CC breast, colon, mouth or skin (especially a melanoma), and angiogenic  
 CC disorders such as retinal neovascularisation. AAS15383-AAS15422  
 CC represent the FAK antisense oligonucleotides of the invention.  
 XX  
 SQ Sequence 20 BP; 5 A; 5 C; 3 G; 7 T; 0 other;  
 XX  
 Query Match 100.0%; Score 20; DB 24; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 0.78;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 TTTCACCGAGATGTCATTC 20  
 Db 1 TTTCACCGAGATGTCATTC 20  
 RESULT 3  
 ABK38943  
 ID ABK38943 standard; cDNA; 550 BP.  
 XX  
 AC ABK38943;  
 XX  
 DT 21-MAY-2002 (first entry)  
 XX  
 DE cDNA encoding lung tumour protein clone R0127.G07.  
 XX  
 KW Lung tumour; cancer; T cell; immune response stimulator;  
 KW cytosstatic; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200204514-A2.  
 XX  
 PD 17-JAN-2002.  
 XX  
 PF 10-JUL-2001; 2001WO-US22058.  
 XX  
 PR 11-JUL-2000; 2000US-0614124.  
 PR 29-AUG-2000; 2000US-0651563.  
 PR 08-SEP-2000; 2000US-0658824.  
 PR 26-SEP-2000; 2000US-0671325.  
 PR 06-OCT-2000; 2000US-0677419.  
 PR 30-OCT-2000; 2000US-0702705.  
 PR 13-DEC-2000; 2000US-0736457.  
 PR 03-MAY-2001; 2001US-0849626.  
 XX  
 PA (CORI-) CORIXA CORP.  
 XX  
 PI Wang T, Watanabe Y, Henderson RA, Johnson JC, Rector MW,  
 PI Marnerakis M, Carter D, Fanger GR, Vedvick TS, Bangur CS;  
 PI McNabb A, Wang A, Fanger N, Switzer A, McNeill PD, Clapper JD;  
 XX  
 DR WPI; 2002-164634/21.  
 XX  
 PT Novel polynucleotide encoding a lung tumour polypeptide useful for  
 PT stimulating and/or expanding T cells specific for a tumour protein -  
 XX  
 PS Claim 1; SEQ ID No 981; 223pp; English.  
 XX  
 CC The invention describes an isolated polynucleotide and polypeptide  
 CC useful for stimulating and/or expanding T cells specific for a tumour  
 CC protein for determining the presence of a cancer in a patient. A  
 CC composition containing the polynucleotide and/or polypeptide is useful  
 CC for treating a lung cancer in a patient. The polypeptide is useful for  
 CC removing tumour cells from a biological sample. The polynucleotide is

CC also useful as probe or primer to detect the level of mRNA encoding a  
CC tumour protein. This sequence encodes a lung tumour associated protein  
CC or protein fragment, described in the method of the invention.  
CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
CX  
SQ Sequence 550 BP; 155 A; 113 C; 121 G; 161 T; 0 other;  
  
Query Match 100.0%; Score 20; DB 24; Length 550;  
Best Local Similarity 100.0%; Pred. No. 1.4;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 TTTCACACAGATGGTCATTC 20  
DB 195 TTTCACACAGATGGTCATTC 214  
  
RESULT 4  
AAC65534/c  
ID AAC65534 standard; DNA; 3791 BP.  
AC  
AC AAC65534;  
XX  
DT 12-FEB-2001 (first entry)  
XX  
DE Human focal adhesion kinase coding sequence.  
XX  
KW Human; focal adhesion kinase; FAK; signal transduction; cancer;  
KW embryonic development disorder; angiogenic disorder; wound healing; ds.  
XX  
OS Homo sapiens.  
XX  
PN US6133031-A.  
XX  
PD 17-OCT-2000.  
XX  
PF 19-AUG-1999; 99US-0377310.  
XX  
PR 19-AUG-1999; 99US-0377310.  
XX  
PA (ISIS-) ISIS PHARM INC.  
XX  
PI Monia BP, Gaarde WA;  
XX  
DR WPI; 2001-006141/01.  
XX  
PT P-PSDB; AAB30327.  
XX  
PT New antisense compounds for inhibiting focal adhesion kinase  
PT expression, especially useful for inhibiting retinal  
PT neovascularization, or for diagnosing and treating e.g. colon cancer -  
XX  
PS Claim 1; Column 31-40; 30pp; English.  
XX  
CC The present invention describes a number of phosphorothioate antisense  
CC sequences to the human focal adhesion kinase (FAK) protein. This protein  
CC is involved in integrin-mediated signal transduction, and is implicated  
CC in cancer, particularly colon, breast and oral tumours, embryonic  
CC development disorders, angiogenic disorders and wound healing. The  
CC antisense sequences can be used in the treatment of all of these.  
XX  
SQ Sequence 3791 BP; 1110 A; 849 C; 924 G; 908 T; 0 other;  
  
Query Match 100.0%; Score 20; DB 22; Length 3791;  
Best Local Similarity 100.0%; Pred. No. 1.9;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 TTTCACACAGATGGTCATTC 20  
DB 3529 TTTCACACAGATGGTCATTC 3510  
  
also useful as probe or primer to detect the level of mRNA encoding a  
tumour protein. This sequence encodes a lung tumour associated protein  
or protein fragment, described in the method of the invention.  
Note: The sequence data for this patent did not form part of the  
printed specification, but was obtained in electronic format directly  
from WIPO at  
ftp.wipo.int/pub/published\_pct\_sequences.  
Sequence 550 BP; 155 A; 113 C; 121 G; 161 T; 0 other;  
  
Query Match 100.0%; Score 20; DB 24; Length 550;  
Best Local Similarity 100.0%; Pred. No. 1.4;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 TTTCACACAGATGGTCATTC 20  
DB 195 TTTCACACAGATGGTCATTC 214  
  
RESULT 5  
AAS15382/c  
ID AAS15382 standard; cDNA; 3791 BP.  
XX  
AC AAS15382;  
XX  
DT 29-JAN-2002 (first entry)  
XX  
DE Human focal adhesion kinase (FAK) cDNA sequence.  
XX  
KW Human; focal adhesion kinase; FAK; melanoma; tumour metastasis;  
KW cancer; angiogenic disorder; retinal neovascularisation; cytostatic;  
KW ophthalmological; ss.  
XX  
OS Homo sapiens.  
XX  
PN US2001034329-A1.  
XX  
PD 25-OCT-2001.  
XX  
PF 09-JAN-2001; 2001US-0757100.  
XX  
PR 19-AUG-1999; 99US-0377310.  
PR 13-JUL-2000; 2000WO-US18999.  
XX  
PA (MONI/) MONIA B P.  
PA (GAAR/) GAARDE W A.  
PA (NERO/) NERO P S.  
XX  
PI Monia BP, Gaarde WA, Nero PS;  
XX  
DR WPI; 2002-010103/01.  
XX  
PT Antisense suppression of Focal Adhesion Kinase expression for the  
PT treatment of cancers of the breast, colon, mouth or skin (especially a  
PT melanoma), and angiogenic disorders e.g. retinal neovascularisation -  
XX  
PS Example 2; Page -; 19pp; English.  
XX  
CC The present invention relates to novel antisense compounds which can be  
CC used for modulating the expression of human focal adhesion kinase (FAK).  
CC The antisense compounds comprise antisense oligonucleotides  
CC (8-30 nucleotides in length) targeted to the 5'-untranslated region,  
CC translational termination region or 3' untranslated region of a nucleic  
CC acid molecule encoding FAK. The antisense oligonucleotides can be used  
CC to inhibit the expression of FAK mRNA. The antisense oligonucleotides  
CC represent potential chemotherapeutic agents in the treatment of melanoma  
CC and the prevention of tumour metastasis. The antisense compounds and  
CC methods are useful for treating diseases associated with overexpression  
CC or constitutive activation of FAK. Such diseases include cancers of the  
CC breast, colon, mouth or skin (especially a melanoma), and angiogenic  
CC disorders such as retinal neovascularisation. The present cDNA sequence  
CC for human FAK is used to design the FAK antisense oligonucleotides  
CC of the invention.  
CC Note: This sequence (Genbank accession number U13616) is not given in  
CC the patent but is acquired by the indexer from Genbank from the  
CC information provided in the patent.  
XX  
SQ Sequence 3791 BP; 1110 A; 849 C; 924 G; 908 T; 0 other;  
  
Query Match 100.0%; Score 20; DB 24; Length 3791;  
Best Local Similarity 100.0%; Pred. No. 1.9;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 TTTCACACAGATGGTCATTC 20  
DB 3529 TTTCACACAGATGGTCATTC 3510  
  
RESULT 6  
ABV29192/c  
ID ABV29192 standard; cDNA; 4052 BP.  
XX



PA (CHIB-) CHIBA PREFECTURE.  
PA (HISM ) HISAMITSU PHARM CO LTD.  
XX Nakagawara A;  
XX WPI; 2001-565584/63.  
XX  
XX Nucleic acids originating in gene expressed in human neuroblastoma,  
PT useful as probe or primer in diagnosing prognosis of human  
PT neuroblastoma, malignancy and susceptibility indicator or tumour marker  
PT for anti-cancer agents -  
XX  
XX Claim 1; Page 2870-2871; 2979pp; Japanese.  
XX  
XX The invention relates to novel genes (AAI93926-AAI97963) expressed in  
CC human neuroblastoma. The nucleic acids are applicable as a probe or  
CC primer in diagnosing the prognosis of human neuroblastoma, malignancy and  
CC susceptibility indicators or tumour markers for anti-cancer agents. The  
CC gene information for diagnosing prognosis is related to factors similar  
CC to that for N-myc and TrkA genes.  
XX  
XX Sequence 707 BP; 200 A; 129 C; 149 G; 224 T; 5 other;  
SQ  
Query Match 92.0%; Score 18.4; DB 22; Length 707;  
Best Local Similarity 95.0%; Pred. No. 9.8;  
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 TTTCACCCAGATGTCATTC 20  
|||||  
Db 431 TTTCACCCAGATGTCATTC 450  
RESULT 9  
ABA93211  
ID ABA93211 standard; DNA; 1311 BP.  
XX  
AC ABA93211;  
XX  
DT 17-APR-2002 (first entry)  
XX  
DE Coriolus versicolor aldo/ketoreductase encoding DNA SEQ ID NO:1.  
XX  
KW Aldo/ketoreductase; white-rot basidiomycete; methylbenzothioephene; AAD;  
KW aryl-alcohol dehydrogenase; enzyme; gene; ds.  
XX  
OS Coriolus versicolor.  
XX  
XX Key Location/Qualifiers  
FH CDS 354..1229  
FT /\*tag= a  
FT /product= "aldo/ketoreductase"  
XX  
PN JP2001321171-A.  
XX  
PD 20-NOV-2001.  
XX  
PF 21-MAR-2000; 2000JP-0078691.  
XX  
PR 09-MAR-2000; 2000JP-0065559.  
XX  
PA (WARI/) WARIISHI H.  
PA (KUBI ) KUBOTA CORP.  
XX  
XX WPI; 2002-158675/21.  
DR P-PSDB; ABB05467.  
XX  
XX A gene encoding a new protein and its use -  
PT  
XX Claim 1; Page 7-8; 12pp; Japanese.  
XX  
XX The present sequence encodes an aldo/ketoreductase (I) isolated from the  
CC white rot basidiomycete, Coriolus versicolor. The present invention also  
CC describes a method for detecting at least one sulphur-containing

CC hydrocarbon compound selected from methylbenzothiophenes in which, in a  
CC process of assaying a compound by the expression of (I), the increase in  
CC the expression of (I) shows that the compound is the objective compound.  
CC (I) and the gene encoding (I) can be used for decomposing noxious  
CC substances, particularly sulphur-containing heterocyclic compounds.  
XX  
SQ Sequence 1311 BP; 340 A; 332 C; 391 G; 248 T; 0 other;  
Query Match 85.0%; Score 17; DB 24; Length 1311;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 TTCAACCCAGATGTCAT 18  
|||||  
Db 357 TTCAACCCAGATGTCAT 373  
RESULT 10  
ABA51579  
ID ABA51579 standard; DNA; 244 BP.  
XX  
AC ABA51579;  
XX  
DT 01-FEB-2002 (first entry)  
XX  
DE Human breast cell single exon nucleic acid probe #10274.  
XX  
KW Human; microarray; single exon probe; gene expression; breast;  
KW disease; cancer; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200157271-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 30-JAN-2001; 2001WO-US00662.  
XX  
PR 04-FEB-2000; 2000US-0180312.  
PR 26-MAY-2000; 2000US-0207456.  
PR 30-JUN-2000; 2000US-0608408.  
PR 03-AUG-2000; 2000US-0632366.  
PR 21-SEP-2000; 2000US-0234687.  
PR 27-SEP-2000; 2000US-0236359.  
PR 04-OCT-2000; 2000GB-0024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
XX Penn SG, Hanzel DK, Chen W, Rank DR;  
PI WPI; 2001-496933/54.  
XX  
XX New spatially-addressable set of single exon nucleic acid probes,  
PT useful for measuring gene expression in sample derived from human  
PT breast, comprises number of single exon nucleic acid probes -  
XX  
XX Claim 4; SEQ ID NO 10274; 327pp + sequence listing; English.  
XX  
XX The invention relates to a spatially-addressable set of single exon  
CC nucleic acid probes for measuring gene expression in a sample derived  
CC from human breast and BT 474 cells. The method involves contacting  
CC the probes with a collection of detectably labelled nucleic acids  
CC derived from mRNA of human breast, and then measuring the label  
CC bound to each probe of the microarray. The probes are useful for  
CC verifying the expression of regions of genomic DNA predicted to  
CC encode proteins. They are useful for gene discovery, and for  
CC determining predisposition and/or prognosing breast disease. Gene  
CC expression analysis is useful for assessing the toxicity of chemical  
CC agents on cells. The microarray of this invention presents a far greater  
CC diversity of probes for measuring gene expression, with far less bias  
CC than expressed sequence tag microarrays. The method is suitable for  
CC rapid production of functional information from genomic sequence. The  
CC present sequence is a single exon nucleic acid probe of the invention.

CC Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 244 BP; 50 A; 59 C; 36 G; 99 T; 0 other;

Query Match 84.0%; Score 16.8; DB 22; Length 244;

Best Local Similarity 90.0%; Pred. No. 56; Mismatches 2; Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Oy 1 TTTCACCGAGTGCATTC 20  
Db 118 TTTCACCGAGTGCATTC 137

RESULT 11  
ABA36555  
ID ABA36555 standard; DNA; 244 BP.

XX AC ABA36555;

XX DT 23-JAN-2002 (first entry)

DE Probe #15021 for gene expression analysis in human heart cell sample.

XX Human; gene expression; heart; microarray; vascular system; probe;

KM cardiovascular disease; hypertension; cardiac arrhythmia;

KM congenital heart disease; ss.

OS Homo sapiens.

XX WO200157274-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00666.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-488899/53.

XX PT Single exon nucleic acid probes for analyzing gene expression in human

XX PT hearts -

XX PS Claim 4; SEQ ID No 15021; 530bp; English.

XX CC The present invention relates to single exon nucleic acid probes for

XX CC measuring human gene expression in a sample derived from human heart. The

XX CC present sequence is one such probe. The probes may be used for

XX CC predicting, measuring and displaying gene expression in samples derived

XX CC from the human heart via microarrays. By measuring gene expression, the

XX CC probes are useful for predicting, diagnosing, grading, staging, the

XX CC monitoring and prognosing diseases of the human heart and vascular system

XX CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and

XX CC congenital heart disease.

XX CC Note: The sequence data for this patent did not form part of the printed

XX CC specification, but was obtained in electronic format directly from WIPO

XX CC at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 244 BP; 50 A; 59 C; 36 G; 99 T; 0 other;

Query Match 84.0%; Score 16.8; DB 22; Length 244;

Best Local Similarity 90.0%; Pred. No. 56;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 TTTCACCGAGTGCATTC 20  
Db 118 TTTCACCGAGTGCATTC 137

RESULT 12  
AAK43716  
ID AAK43716 standard; DNA; 244 BP.

XX AC AAK43716;

XX DT 06-NOV-2001 (first entry)

DE Human bone marrow expressed single exon probe SEQ ID NO: 18273.

XX Human; bone marrow expressed exon; gene expression analysis; probe;

KM microarray; cancer; leukemia; lymphoma; myeloma; ss.

XX OS Homo sapiens.

XX PN WO200157276-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00668.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-488900/53.

XX PT Human genome-derived single exon nucleic acid probes useful for

XX PT analyzing gene expression in human bone marrow -

XX PS Example 4; SEQ ID NO: 18273; 658bp + Sequence Listing; English.

XX CC The present invention provides a number of single exon nucleic acid

XX CC probes which are derived from genomic sequences expressed in the human

XX CC bone marrow. They can be used to measure gene expression in bone marrow

XX CC samples, which may enable the improved diagnosis and treatment of cancers

XX CC such as lymphoma, leukemia and myeloma. The present sequence is one of

XX CC the probes of the invention.

XX SQ Sequence 244 BP; 50 A; 59 C; 36 G; 99 T; 0 other;

Query Match 84.0%; Score 16.8; DB 22; Length 244;

Best Local Similarity 90.0%; Pred. No. 56;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 TTTCACCGAGTGCATTC 20

Db 118 TTTCACCGAGTGCATTC 137

RESULT 13

ABA46482

ID ABA46482 standard; DNA; 400 BP.

XX AC ABA46482;

XX DT 01-FEB-2002 (first entry)



```
DE Human breast cell single exon nucleic acid probe #5177.
XX Human; microarray; single exon probe; gene expression; breast;
KW disease; cancer; ss.
XX Homo sapiens.
XX OS
XX PN WO200157271-A2.
XX PD
XX PF 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US00662.
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX PF WI; 2001-496933/54.
XX PT New spatially-addressable set of single exon nucleic acid probes,
XX PT useful for measuring gene expression in sample derived from human
XX PT breast, comprises number of single exon nucleic acid probes -
XX PS Claim 1; SEQ ID NO 5177; 327pp + sequence listing; English.
XX CC The invention relates to a spatially-addressable set of single exon
XX CC nucleic acid probes for measuring gene expression in a sample derived
XX CC from human breast and BT 474 cells. The method involves contacting
XX CC the probes with a collection of detectably labelled nucleic acids
XX CC derived from mRNA of human breast, and then measuring the label
XX CC bound to each probe of the microarray. The probes are useful for
XX CC verifying the expression of regions of genomic DNA predicted to
XX CC encode proteins. They are useful for gene discovery, and for
XX CC determining predisposition and/or prognosing breast disease. Gene
XX CC expression analysis is useful for assessing the toxicity of chemical
XX CC agents on cells. The microarray of this invention presents a far greater
XX CC diversity of probes for measuring gene expression, with far less bias
XX CC than expressed sequence tag microarrays. The method is suitable for
XX CC rapid production of functional information from genomic sequence. The
XX CC present sequence is a single exon nucleic acid probe of the invention.
XX CC Note: The sequence data for this patent did not form part of the
XX CC printed specification, but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 400 BP; 94 A; 97 C; 84 G; 125 T; 0 other;

Query Match 84.0%; Score 16.8; DB 22; Length 400;
Best Local Similarity 90.0%; Pred. No. 61;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTTCACACGATGGTCATTC 20
Db 344 TTTCACACGATGGTCATTC 363

RESULT 14
ABA26688
ID ABA26688 standard; DNA; 400 BP.
XX AC ABA26688;
XX XX
XX DT 23-JAN-2002 (first entry)
XX DE Probe #5154 for gene expression analysis in human heart cell sample.
XX

KW Human; gene expression; heart; microarray; vascular system; probe;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
XX congenital heart disease; ss.
XX OS
XX PN WO200157274-A2.
XX PD
XX PF 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US00666.
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX PF WI; 2001-488899/53.
XX PT Single exon nucleic acid probes for analyzing gene expression in human
XX PT hearts -
XX PS Claim 1; SEQ ID NO 5154; 530pp; English.
XX CC The present invention relates to single exon nucleic acid probes for
XX CC measuring human gene expression in a sample derived from human heart. The
XX CC present sequence is one such probe. The probes may be used for
XX CC predicting, measuring and displaying gene expression in samples derived
XX CC from the human heart via microarrays. By measuring gene expression, the
XX CC probes are useful for predicting, diagnosing, grading, staging,
XX CC monitoring and prognosing diseases of the human heart and vascular system
XX CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
XX CC congenital heart disease.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 400 BP; 94 A; 97 C; 84 G; 125 T; 0 other;

Query Match 84.0%; Score 16.8; DB 22; Length 400;
Best Local Similarity 90.0%; Pred. No. 61;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTTCACACGATGGTCATTC 20
Db 344 TTTCACACGATGGTCATTC 363

RESULT 15
AAK30708
ID AAK30708 standard; DNA; 400 BP.
XX AC AAK30708;
XX XX
XX DT 06-NOV-2001 (first entry)
XX DE Human bone marrow expressed single exon probe SEQ ID NO: 5265.
XX XX
XX KW Human; bone marrow expressed exon; gene expression analysis; probe;
XX KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX OS
XX PN WO200157276-A2.
XX PD 09-AUG-2001.
```

```

XX 30-JAN-2001; 2001WO-US00668.
PF
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488900/53.
XX
DR WPI; 2001-488900/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
XX Example 4; SEQ ID NO: 5265; 658bp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukemia and myeloma. The present sequence is one of
CC the probes of the invention.
XX
SQ Sequence 400 BP; 94 A; 97 C; 84 G; 125 T; 0 other;
Query Match 84.0%; Score 16.8; DB 22; Length 400;
Best Local Similarity 90.0%; Pred. No. 61;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 TTTCACCGAGTGCATTC 20
Db 344 TTTCACCGAGTGCATTC 363

```

Search completed: January 12, 2003, 09:51:21  
 Job time : 255 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2003, 09:46:10 ; Search time 47 Seconds  
(without alignments)  
130.501 Million cell updates/sec

Title: US-09-757-100B-18

Perfect score: 20

Sequence: 1 ttccaaccagatgggtcattc 20

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*

- 1: /cgn2\_6/prodata/1/ina/5A COMB.seq:\*
- 2: /cgn2\_6/prodata/1/ina/5B COMB.seq:\*
- 3: /cgn2\_6/prodata/1/ina/6A COMB.seq:\*
- 4: /cgn2\_6/prodata/1/ina/6B COMB.seq:\*
- 5: /cgn2\_6/prodata/1/ina/PCITUS COMB.seq:\*
- 6: /cgn2\_6/prodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	20	100.0	20	3	US-09-377-310-18
2	20	100.0	3791	3	US-09-377-310-1
3	16.8	84.0	4465	1	US-08-180-195-1
4	16.8	84.0	4465	1	US-08-477-329-1
5	16.8	84.0	4465	2	US-08-475-458-1
6	16.8	84.0	4465	3	US-08-980-400-1
7	16.8	84.0	4465	4	US-09-583-459A-1
8	16.8	84.0	4465	4	US-09-583-210-1
9	16.8	84.0	4465	4	US-09-583-449A-1
10	16.8	84.0	4465	4	US-09-435-059-1
11	16.8	84.0	5427	1	US-08-168-917-1
12	16.8	84.0	5427	2	US-08-460-510-1
13	16.8	84.0	5427	2	US-08-460-490-1
14	16.8	84.0	5427	3	US-08-462-728-3
15	16.8	84.0	5427	4	US-08-461-917-3
16	16.8	84.0	5427	5	PCT-US92-00730-1
17	16.8	84.0	5427	5	PCT-US92-00862-1
18	15.8	79.0	7902	4	US-08-961-527-112
19	15.8	79.0	10223	4	US-08-961-527-73
20	15.2	76.0	3233	3	US-08-755-587-43
21	15.2	76.0	5313	4	US-09-329-920-1
22	15.2	76.0	6169	2	US-08-875-154-2
23	15.2	76.0	11385	2	US-08-639-501-1
24	15.2	76.0	11385	3	US-09-044-946-1
25	15.2	76.0	11385	3	US-09-044-908-1
26	15.2	76.0	87563	4	US-09-453-702B-57
27	15	75.0	15	3	US-09-377-310-38

```

28 14.8 74.0 969 1 US-08-564-596B-4 Sequence 4, Appli
29 14.8 74.0 969 1 US-08-721-798A-1 Sequence 1, Appli
30 14.8 74.0 3266 4 US-08-485-511A-3 Sequence 3, Appli
C 31 14.8 74.0 8136 4 US-08-961-527-75 Sequence 75, Appl
C 32 14.8 74.0 246240 2 US-08-724-394A-20 Sequence 20, Appl
C 33 14.8 74.0 246240 2 US-08-724-394A-21 Sequence 21, Appl
C 34 14.8 74.0 246240 2 US-08-724-394A-22 Sequence 22, Appl
C 35 14.8 72.0 38 4 US-09-233-086-53 Sequence 53, Appl
36 14.4 72.0 1001 4 US-09-641-638-91 Sequence 91, Appl
37 14.4 72.0 1713 3 US-08-467-948A-1 Sequence 1, Appli
38 14.4 72.0 1713 3 US-08-467-947A-1 Sequence 1, Appli
39 14.4 72.0 1960 4 US-09-178-252-19 Sequence 19, Appl
C 40 14.4 72.0 2580 2 US-08-887-798-1 Sequence 1, Appli
41 14.4 72.0 3489 4 US-09-178-252-20 Sequence 20, Appl
42 14.2 71.0 20 3 US-09-357-071-39 Sequence 39, Appl
43 14.2 71.0 150 2 US-08-250-346-1 Sequence 1, Appli
44 14.2 71.0 267 4 US-09-696-569-7 Sequence 7, Appli
45 14.2 71.0 536 4 US-09-347-819-7 Sequence 7, Appli

```

#### ALIGNMENTS

```

RESULT 1
US-09-377-310-18
; Sequence 18, Application US/09377310B
; Patent No. 6133031
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; FILE OF INVENTION: Expression
; FILE REFERENCE: ISPH-0389
; CURRENT APPLICATION NUMBER: US/09/377,310B
; CURRENT FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 18
; LENGTH: 20
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: antisense sequence
US-09-377-310-18

```

```

Query Match 100.0%; Score 20; DB 3; Length 20;
Best Local Similarity 100.0%; Pred No. 0.14;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 TTTCACCAAGATGGTCATTC 20
|||||
DB 1 TTTCACCAAGATGGTCATTC 20
|||||

```

```

RESULT 2
US-09-377-310-1/c
; Sequence 1, Application US/09377310B
; Patent No. 6133031
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; FILE OF INVENTION: Expression
; FILE REFERENCE: ISPH-0389
; CURRENT APPLICATION NUMBER: US/09/377,310B
; CURRENT FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3791
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

```

NAME/KEY: CDS  
LOCATION: (233)..(3391)  
PUBLICATION INFORMATION:  
JOURNAL: DNA  
VOLUME: 12  
ISSUE: 9  
PAGES: 823-830  
DATE: 1993-11  
DATABASE ACCESSION NUMBER: L13616/Genbank  
DATABASE ENTRY DATE: 1995-01-02  
US-09-377-310-1

Query Match 100.0%; Score 20; DB 3; Length 3791;  
Best Local Similarity 100.0%; Pred. NO. 0.28;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 TTTCACCGAGTGGTCATTC 20  
Db 3529 TTTCACCGAGTGGTCATTC 3510

## RESULT 3

US-08-180-195-1/c  
Sequence 1, Application US/08180195  
Patent No. 5567584  
GENERAL INFORMATION:  
APPLICANT: Sledziewski Ph.D., Andrzej Z  
APPLICANT: Bell, Lillian A.  
APPLICANT: Kindsvogel Ph.D., Wayne R.  
TITLE OF INVENTION: METHODS OF PRODUCING SECRETED RECEPTOR ANALOGS  
TITLE OF INVENTION: AND BIOLOGICALLY ACTIVE DIMERIZED POLYPEPTIDE  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seed and Berry  
STREET: 6300 Columbia Center  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.24  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/180.195  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/634.510  
FILING DATE:  
APPLICATION NUMBER: US 07/146.877  
FILING DATE: 22-JAN-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/347.291  
FILING DATE: 02-MAY-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki J.D., David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 390008.446C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 206-622-4900  
TELEFAX: 206-682-6031  
TELEX: 3723836  
INFORMATION FOR SEQ ID NO. 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4465 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: N

ANTI-SENSE: N  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
DEVELOPMENTAL STAGE: Adult  
TISSUE TYPE: Skin  
CELL TYPE: fibroblasts  
IMMEDIATE SOURCE:  
CLONE: pr-rx1  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 354..3671  
OTHER INFORMATION:  
US-08-180-195-1

Query Match 84.0%; Score 16.8; DB 1; Length 4465;  
Best Local Similarity 90.0%; Pred. NO. 13;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 TTTCACCGAGTGGTCATTC 20  
Db 1263 TTTCACCGAGTGGTCATTC 1244

## RESULT 4

US-08-477-329-1/c  
Sequence 1, Application US/08477329  
Patent No. 5750375  
GENERAL INFORMATION:  
APPLICANT: Sledziewski Ph.D., Andrzej Z  
APPLICANT: Bell, Lillian A.  
APPLICANT: Kindsvogel Ph.D., Wayne R.  
TITLE OF INVENTION: METHODS OF PRODUCING SECRETED RECEPTOR ANALOGS AND BIOLOGICAL  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seed and Berry  
STREET: 6300 Columbia Center, 701 Fifth Avenue  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98104-7092  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.24  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477.329  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Maki, David J.  
REGISTRATION NUMBER: 31,392  
REFERENCE/DOCKET NUMBER: 390008.446C6  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 206-622-4900  
TELEFAX: 206-682-6031  
TELEX: 3723836  
INFORMATION FOR SEQ ID NO. 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 4465 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: N  
ANTI-SENSE: N  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
DEVELOPMENTAL STAGE: Adult  
TISSUE TYPE: Skin  
CELL TYPE: fibroblasts  
IMMEDIATE SOURCE:  
CLONE: pr-rx1

;  
; FEATURE: CDS  
; NAME/KEY: 354...3671  
; LOCATION: 354...3671  
; OTHER INFORMATION:  
US-08-477-329-1

Query Match 84.0%; Score 16.8; DB 1; Length 4465;

Best Local Similarity 90.0%; Pred. No. 13;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTTCACACAGATGTCATTC 20

Db 1263 TTTCATCCTGATGTCATTC 1244

## RESULT 5

US-08-475-458-1/c

; Sequence 1, Application US/08475458

; Patent No. 5843725

; GENERAL INFORMATION:

; APPLICANT: Siedziwski Ph.D., Andrzej Z

; APPLICANT: Bell, Lillian A.

; APPLICANT: Kindsvogel Ph.D., Wayne R.

; TITLE OF INVENTION: METHODS OF PRODUCING SECRETED RECEPTOR ANALOGS AND BIOLOGICAL

; NUMBER OF SEQUENCES: 36

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Seed and Berry

; STREET: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: WA

; COUNTRY: USA

; ZIP: 98104-7092

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.24

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/475,458

; FILING DATE: 06-JUN-1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Maki, David J.

; REGISTRATION NUMBER: 31,392

; REFERENCE/DOCKET NUMBER: 990008.446D5

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 206-622-4900

; TELEFAX: 206-682-6031

; TELEX: 3723836

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 4465 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; HYPOTHETICAL: N

; ANTI-SENSE: N

; ORIGINAL SOURCE:

; ORGANISM: Homo sapiens

; DEVELOPMENTAL STAGE: Adult

; TISSUE TYPE: Skin

; CELL TYPE: fibroblasts

; IMMEDIATE SOURCE:

; CLONE: PR-rX1

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 354...3671

; OTHER INFORMATION:

US-08-475-458-1

Query Match

Best Local Similarity 84.0%; Score 16.8; DB 2; Length 4465;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTTCACACAGATGTCATTC 20

Db 1263 TTTCATCCTGATGTCATTC 1244

## RESULT 6

US-08-980-400-1/c

; Sequence 1, Application US/08980400

; Patent No. 6018026

; GENERAL INFORMATION:

; APPLICANT: Siedziwski Ph.D., Andrzej Z

; APPLICANT: Bell, Lillian A.

; APPLICANT: Kindsvogel Ph.D., Wayne R.

; TITLE OF INVENTION: METHODS OF PRODUCING SECRETED RECEPTOR ANALOGS AND BIOLOGICAL

; NUMBER OF SEQUENCES: 36

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Seed and Berry

; STREET: 6300 Columbia Center, 701 Fifth Avenue

; CITY: Seattle

; STATE: WA

; COUNTRY: USA

; ZIP: 98104-7092

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.24

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/980,400

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/477,329

; FILING DATE: 07-JUN-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Maki, David J.

; REGISTRATION NUMBER: 31,392

; REFERENCE/DOCKET NUMBER: 990008.446C6

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 206-622-4900

; TELEFAX: 206-682-6031

; TELEX: 3723836

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 4465 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; HYPOTHETICAL: N

; ANTI-SENSE: N

; ORIGINAL SOURCE:

; ORGANISM: Homo sapiens

; DEVELOPMENTAL STAGE: Adult

; TISSUE TYPE: Skin

; CELL TYPE: fibroblasts

; IMMEDIATE SOURCE:

; CLONE: PR-rX1

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 354...3671

; OTHER INFORMATION:

US-08-980-400-1

Query Match 84.0%; Score 16.8; DB 3; Length 4465;

Best Local Similarity 90.0%; Pred. No. 13;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTTCACACAGATGTCATTC 20

Db 1263 TTTCATCCTGATGTCATTC 1244

## RESULT 7

US-09-583-459A-1/c

Sequence 1, Application US/09583459A

Patent No. 6291212

## GENERAL INFORMATION:

APPLICANT: Siegfiewski Ph.D., Andrzej Z

APPLICANT: Bell, Lillian A.

APPLICANT: Kindavogel Ph.D., Wayne R.

TITLE OF INVENTION: METHODS OF PRODUCING SECRETED RECEPTOR ANALOGS

TITLE OF INVENTION: AND BIOLOGICALLY ACTIVE DIMERIZED POLYPEPTIDE

TITLE OF INVENTION: FUSIONS

NUMBER OF SEQUENCES: 36

CORRESPONDENCE ADDRESS:

ADDRESSEE: Seed and Berry

STREET: 6300 Columbia Center

CITY: Seattle

STATE: WA

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.24

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/583,459A

FILING DATE: 30-MAY-2000

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/634,510

FILING DATE: 27-DEC-1990

APPLICATION NUMBER: US 07/146,877

FILING DATE: 22-JAN-1988

APPLICATION NUMBER: US 07/347,291

FILING DATE: 02-MAY-1989

ATTORNEY/AGENT INFORMATION:

NAME: Maki J.D., David J.

REGISTRATION NUMBER: 31,392

REFERENCE/DOCKET NUMBER: 990008.446C3

TELECOMMUNICATION INFORMATION:

TELEPHONE: 206-622-4900

TELEFAX: 206-682-6031

TELEX: 3723836

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 4465 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: CDNA

HYPOTHETICAL: N

ANTI-SENSE: N

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

DEVELOPMENTAL STAGE: Adult

TISSUE TYPE: Skin

CELL TYPE: fibroblasts

IMMEDIATE SOURCE:

CLONE: pr-ix1

FEATURE:

NAME/KEY: CDS

LOCATION: 354..3671

OTHER INFORMATION:

US-09-583-459A-1

Query Match 84.0%; Score 16.8; DB 4; Length 4465;

Best Local Similarity 90.0%; Pred.No. 13;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTTCACCGATGTCATTC 20

|||||

Db 1263 TTTCATCCTGATGTCATTC 1244

## RESULT 8

US-09-583-210-1/c

Sequence 1, Application US/09583210

Patent No. 6291646

## GENERAL INFORMATION:

APPLICANT: Siegfiewski Ph.D., Andrzej Z

APPLICANT: Bell, Lillian A.

APPLICANT: Kindavogel Ph.D., Wayne R.

TITLE OF INVENTION: METHODS OF PRODUCING SECRETED RECEPTOR ANALOGS

TITLE OF INVENTION: AND BIOLOGICALLY ACTIVE DIMERIZED POLYPEPTIDE

TITLE OF INVENTION: FUSIONS

NUMBER OF SEQUENCES: 36

CORRESPONDENCE ADDRESS:

ADDRESSEE: Seed and Berry

STREET: 6300 Columbia Center

CITY: Seattle

STATE: WA

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patentin Release #1.24

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/583,210

FILING DATE: 30-MAY-2000

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/634,510

FILING DATE: 27-DEC-1990

APPLICATION NUMBER: US 07/146,877

FILING DATE: 22-JAN-1988

APPLICATION NUMBER: US 07/347,291

FILING DATE: 02-MAY-1989

ATTORNEY/AGENT INFORMATION:

NAME: Maki J.D., David J.

REGISTRATION NUMBER: 31,392

REFERENCE/DOCKET NUMBER: 990008.446C3

TELECOMMUNICATION INFORMATION:

TELEPHONE: 206-622-4900

TELEFAX: 206-682-6031

TELEX: 3723836

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 4465 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: CDNA

HYPOTHETICAL: N

ANTI-SENSE: N

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

DEVELOPMENTAL STAGE: Adult

TISSUE TYPE: Skin

CELL TYPE: fibroblasts

IMMEDIATE SOURCE:

CLONE: pr-ix1

FEATURE:

NAME/KEY: CDS

LOCATION: 354..3671

OTHER INFORMATION:

US-09-583-210-1

Query Match 84.0%; Score 16.8; DB 4; Length 4465;

Best Local Similarity 90.0%; Pred.No. 13;

Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTTCACCGATGTCATTC 20

|||||

Db 1263 TTTCATCCTGATGGTCATTC 1244  
|||||

RESULT 9  
US-09-583-449A-1/c  
; Sequence 1, Application US/09583449A  
; Patent No. 6300099  
; GENERAL INFORMATION:  
; APPLICANT: Sledziewski Ph.D., Andrzej Z  
; APPLICANT: Bell, Lillian A.  
; APPLICANT: Kindsvogel Ph.D., Wayne R.  
; TITLE OF INVENTION: METHODS OF PRODUCING SECRETED RECEPTOR ANALOGS  
; TITLE OF INVENTION: AND BIOLOGICALLY ACTIVE DIMERIZED POLYPEPTIDE  
; TITLE OF INVENTION: FUSIONS  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seed and Berry  
; STREET: 6300 Columbia Center  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.24  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/583,449A  
; FILING DATE: 30-MAY-2000  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/634,510  
; FILING DATE: 27-DEC-1990  
; APPLICATION NUMBER: US 07/146,877  
; FILING DATE: 22-JAN-1988  
; APPLICATION NUMBER: US 07/347,291  
; FILING DATE: 02-MAY-1989  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maki J.D., David J.  
; REGISTRATION NUMBER: 31,392  
; REFERENCE/DOCKET NUMBER: 990008.446C3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 206-622-4900  
; TELEFAX: 206-682-6031  
; TELEX: 3723836  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4465 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: N  
; ANTI-SENSE: N  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; DEVELOPMENTAL STAGE: Adult  
; TISSUE TYPE: Skin  
; CELL TYPE: fibroblasts  
; IMMEDIATE SOURCE:  
; CLONE: pr-rx1  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 354..3671  
; OTHER INFORMATION:  
US-09-583-449A-1  
Query Match 84.0%; Score 16.8; DB 4; Length 4465;  
Best Local Similarity 90.0%; Pred. No. 13;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTTCACCAAGATGGTCATTC 20  
|||||  
Db 1263 TTTCATCCTGATGGTCATTC 1244

RESULT 10  
US-09-435-059-1/c  
; Sequence 1, Application US/09435059  
; Patent No. 6323323  
; GENERAL INFORMATION:  
; APPLICANT: Sledziewski Ph.D., Andrzej Z  
; APPLICANT: Bell, Lillian A.  
; APPLICANT: Kindsvogel Ph.D., Wayne R.  
; TITLE OF INVENTION: METHODS OF PRODUCING SECRETED RECEPTOR ANALOGS AND BIOLOGICAL  
; TITLE OF INVENTION: FUSIONS  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seed and Berry  
; STREET: 6300 Columbia Center, 701 Fifth Avenue  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.24  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/435,059  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/477,329  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Maki, David J.  
; REGISTRATION NUMBER: 31,392  
; REFERENCE/DOCKET NUMBER: 990008.446C6  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 206-622-4900  
; TELEFAX: 206-682-6031  
; TELEX: 3723836  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4465 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: N  
; ANTI-SENSE: N  
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
; DEVELOPMENTAL STAGE: Adult  
; TISSUE TYPE: Skin  
; CELL TYPE: fibroblasts  
; IMMEDIATE SOURCE:  
; CLONE: pr-rx1  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: 354..3671  
; OTHER INFORMATION:  
US-09-435-059-1  
Query Match 84.0%; Score 16.8; DB 4; Length 4465;  
Best Local Similarity 90.0%; Pred. No. 13;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTTCACCAAGATGGTCATTC 20  
|||||  
Db 1263 TTTCATCCTGATGGTCATTC 1244

RESULT 11  
US-08-168-917-1/c  
Sequence 1, Application US/08168917  
Patent No. 5686572  
GENERAL INFORMATION:  
APPLICANT: Wolf, David  
APPLICANT: Tomlinson, James E.  
APPLICANT: Fretto, Larry J.  
APPLICANT: Giese, Neill A.  
APPLICANT: Escobedo, Jaime A.  
APPLICANT: Williams, Lewis T.  
TITLE OF INVENTION: DOMAINS OF EXTRACELLULAR REGION OF HUMAN  
PLATELET-DERIVED GROWTH FACTOR RECEPTOR POLYPEPTIDES  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TOWNSEND and TOWNSEND  
STREET: Stewart Street Tower, 20th Floor \ One Market  
CITY: San Francisco  
STATE: California  
COUNTRY: US  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
FILING DATE:  
APPLICATION NUMBER: US/08/168,917  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/650,793  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Ching, Edwin P.  
REGISTRATION NUMBER: 34,090  
REFERENCE/DOCKET NUMBER: 12418-14  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 326-2400  
TELEFAX: (415) 326-2400  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5427 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA to mRNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo Sapiens  
STRAIN: lambda gt10  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 187..3504  
US-08-168-917-1

Query Match 84.0%; Score 16.8; DB 1; Length 5427;  
Best Local Similarity 90.0%; Pred. No. 13;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTTCACCGATGTCATTC 20  
DB 1096 TTTCATCTGATGTCATTC 1077

RESULT 12  
US-08-460-510-1/c  
Sequence 1, Application US/08460510  
Patent No. 5872218  
GENERAL INFORMATION:  
APPLICANT: Wolf, David

APPLICANT: Tomlinson, James E.  
APPLICANT: Fretto, Larry J.  
APPLICANT: Giese, Neill A.  
APPLICANT: Escobedo, Jaime A.  
APPLICANT: Williams, Lewis T.  
TITLE OF INVENTION: DOMAINS OF EXTRACELLULAR REGION OF HUMAN  
PLATELET-DERIVED GROWTH FACTOR RECEPTOR POLYPEPTIDES  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TOWNSEND and TOWNSEND and CREW  
STREET: One Market Plaza, Stewart Street Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: US  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
FILING DATE: 02-JUN-1995  
APPLICATION NUMBER: US/08/460,510  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Dow, Karen B.  
REGISTRATION NUMBER: 29,684  
REFERENCE/DOCKET NUMBER: 012418-001430  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 326-2400  
TELEFAX: (415) 326-2400  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5427 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA to mRNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo Sapiens  
STRAIN: lambda gt10  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 187..3504  
US-08-460-510-1

Query Match 84.0%; Score 16.8; DB 2; Length 5427;  
Best Local Similarity 90.0%; Pred. No. 13;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTTCACCGATGTCATTC 20  
DB 1096 TTTCATCTGATGTCATTC 1077

RESULT 13  
US-08-460-490-1/c  
Sequence 1, Application US/08460490  
Patent No. 5891652  
GENERAL INFORMATION:  
APPLICANT: Wolf, David  
APPLICANT: Tomlinson, James E.  
APPLICANT: Fretto, Larry J.  
APPLICANT: Giese, Neill A.  
APPLICANT: Escobedo, Jaime A.  
APPLICANT: Williams, Lewis T.  
TITLE OF INVENTION: DOMAINS OF EXTRACELLULAR REGION OF HUMAN  
PLATELET-DERIVED GROWTH FACTOR RECEPTOR POLYPEPTIDES  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TOWNSEND and TOWNSEND and CREW



STREET: One Market Plaza, Steuart Street Tower, Suite 2000  
CITY: San Francisco  
STATE: California  
COUNTRY: US  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460,490  
FILING DATE: 02-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Dow, Karen B.  
REGISTRATION NUMBER: 29,684  
REFERENCE/DOCKET NUMBER: 012418-001420  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 326-2400  
TELEFAX: (415) 326-2422  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5427 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo Sapiens  
STRAIN: lambda gt10  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 187..3504  
US-08-460-490-1

Query Match 84.0%; Score 16.8; DB 2; Length 5427;  
Best Local Similarity 90.0%; Pred. No. 13;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTTCACACGATGGTCATTC 20  
||||| ||||| ||||| ||||| |||||  
Db 1096 TTTCATCTCGATGGTCATTC 1077

RESULT 14  
US-08-462-728-3/c  
Sequence 3, Application US/08462728  
Patent No. 6043211  
GENERAL INFORMATION:  
APPLICANT: WILLIAMS, Lewis T.  
APPLICANT: ESCOBEDO, Jaime A.  
TITLE OF INVENTION: PLATELET-DERIVED GROWTH FACTOR RECEPTORS  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew  
STREET: One Market, Steuart Street Tower, 20th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/462,728  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/309,322  
FILING DATE: 10-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/151,414  
FILING DATE: 02-FEB-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Dow, Karen B.  
REGISTRATION NUMBER: 29,684  
REFERENCE/DOCKET NUMBER: 2307K-267-2-2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/326-2400  
TELEFAX: 415/326-2422  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5427 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 187..3507  
US-08-462-728-3

Query Match 84.0%; Score 16.8; DB 3; Length 5427;  
Best Local Similarity 90.0%; Pred. No. 13;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTTCACACGATGGTCATTC 20  
||||| ||||| ||||| ||||| |||||  
Db 1096 TTTCATCTCGATGGTCATTC 1077

RESULT 15  
US-08-461-917-3/c  
Sequence 3, Application US/08461917  
Patent No. 6372438  
GENERAL INFORMATION:  
APPLICANT: WILLIAMS, Lewis T.  
APPLICANT: ESCOBEDO, Jaime A.  
TITLE OF INVENTION: PLATELET-DERIVED GROWTH FACTOR RECEPTORS  
NUMBER OF SEQUENCES: 14  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew  
STREET: One Market, Steuart Street Tower, 20th Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/461,917  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/309,322  
FILING DATE: 10-FEB-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/151,414  
FILING DATE: 02-FEB-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Dow, Karen B.  
REGISTRATION NUMBER: 29,684  
REFERENCE/DOCKET NUMBER: 2307K-267-2-4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/326-2400  
TELEFAX: 415/326-2422  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:

```

; LENGTH: 5427 base pairs
;
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 187..3507
; US-08-461-917-3
    
```

```

Query Match      84.0%; Score:16.8; DB 4; Length 5427;
Best Local Similarity 90.0%; Pred. No.13;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
    
```

```

Qy 1 TTTCACCGATGTCATTC 20
    ||||| ||||| |||||
Db 1096 TTTCATCCTGATGTCATTC 1077
    
```

Search completed: January 12, 2003, 11:11:01  
 Job time : 50 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2003, 10:38:40 ; Search time 56 Seconds  
(without alignments)  
157.180 Million cell updates/sec

Title: US-09-757-100B-18

Perfect score: 20

Sequence: 1 ttccaaccagatggtcattc 20

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 389086 seqs, 220051671 residues

Total number of hits satisfying chosen parameters: 778172

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_5/prodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_5/prodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/prodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/prodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/prodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/prodata/2/pubpna/PCTU5\_PUBCOMB.seq:\*
- 7: /cgn2\_6/prodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/prodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/prodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 10: /cgn2\_6/prodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/prodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 12: /cgn2\_6/prodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 13: /cgn2\_6/prodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/prodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	20	100.0	20	10	US-09-757-100B-18
2	20	100.0	550	9	US-09-736-457-981
3	20	100.0	550	9	US-09-902-941-981
4	20	100.0	550	9	US-09-849-626-981
5	20	100.0	3791	10	US-09-757-100B-1
6	16.8	84.0	244	10	US-09-864-761-21875
7	16.8	84.0	400	10	US-09-864-761-5154
8	16.8	84.0	528	10	US-09-864-761-16204
9	16.8	84.0	1683	10	US-09-866-510-23
10	16.8	84.0	3321	10	US-09-866-510-13
11	16.8	84.0	3321	10	US-09-866-510-15
12	16.8	84.0	3321	10	US-09-866-510-17
13	16.8	84.0	3321	10	US-09-866-510-19
14	16.8	84.0	3321	10	US-09-866-510-21
15	16.8	84.0	4465	9	US-09-955-363-1
16	16.8	84.0	11036	10	US-09-070-927A-464
17	16.8	84.0	32191	10	US-09-764-864-1678
18	15.8	79.0	476	10	US-09-867-701-6708
19	15.8	79.0	575	10	US-09-974-300-5353

C 20	15.8	79.0	1020	10	US-09-815-242-9137	Sequence 9137, Ap
C 21	15.8	79.0	1020	10	US-09-815-242-9537	Sequence 9537, Ap
C 22	15.8	79.0	368004	10	US-09-949-654-3	Sequence 3, Appli
C 23	15.4	77.0	1496	9	US-09-938-842A-3191	Sequence 3191, Ap
C 24	15.2	76.0	321	9	US-09-738-626-242	Sequence 242, App
C 25	15.2	76.0	816	10	US-09-974-300-2088	Sequence 2088, Ap
C 26	15.2	76.0	1134	10	US-09-974-300-180	Sequence 180, App
C 27	15.2	76.0	1413	10	US-09-822-830A-469	Sequence 469, App
C 28	15.2	76.0	9279	10	US-09-801-368-171	Sequence 171, App
C 29	15.2	76.0	3309400	9	US-09-738-626-1	Sequence 1, Appli
C 30	15	75.0	15	10	US-09-757-100B-38	Sequence 38, Appl
C 31	14.8	74.0	530	9	US-09-510-662A-28	Sequence 28, Appl
C 32	14.8	74.0	530	10	US-09-778-320-28	Sequence 28, Appl
C 33	14.8	74.0	530	10	US-09-910-689-28	Sequence 28, Appl
C 34	14.8	74.0	530	12	US-10-010-742-28	Sequence 28, Appl
C 35	14.8	74.0	702	9	US-09-938-842A-975	Sequence 975, App
C 36	14.8	74.0	910	9	US-10-114-893-104	Sequence 104, App
C 37	14.8	74.0	3779	10	US-09-925-300-712	Sequence 712, App
C 38	14.8	74.0	4187	10	US-09-954-456-2112	Sequence 2112, Ap
C 39	14.8	74.0	4187	10	US-09-880-107-1609	Sequence 1609, Ap
C 40	14.8	74.0	41104	10	US-09-816-685-3	Sequence 3, Appli
C 41	14.4	72.0	154	10	US-09-864-761-21937	Sequence 21937, A
C 42	14.4	72.0	369	10	US-09-864-761-3281	Sequence 3281, Ap
C 43	14.4	72.0	454	10	US-09-864-761-14297	Sequence 14297, A
C 44	14.4	72.0	561	10	US-09-864-761-5203	Sequence 5203, Ap
C 45	14.4	72.0	585	10	US-09-747-835A-37	Sequence 37, Appl

#### ALIGNMENTS

RESULT 1  
US-09-757-100B-18  
; Sequence 18, Application US/09757100B  
; Patent No. US20010034329A1  
; GENERAL INFORMATION:  
; APPLICANT: Monia, Brett P.  
; APPLICANT: Gaarde, William A.  
; APPLICANT: Nero, Pamela S.  
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase  
; TITLE OF INVENTION: Expression  
; FILE REFERENCE: ISPH-0533  
; CURRENT APPLICATION NUMBER: US/09/757,100B  
; CURRENT FILING DATE: 2001-03-15  
; PRIOR APPLICATION NUMBER: 09/377,310  
; PRIOR FILING DATE: 1999-08-19  
; PRIOR APPLICATION NUMBER: PCT/US00/18999  
; PRIOR FILING DATE: 2000-07-13  
; NUMBER OF SEQ ID NOS: 44  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 18  
; LENGTH: 20  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: antisense sequence  
US-09-757-100B-18

Query Match 100.0%; Score 20; DB 10; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.13; Mismatches 0; Indels 0; Gaps 0;  
Matches 20; Conservative 0

QY 1 TTTCACACAGATGGTCATTC 20  
|||||  
Db 1 TTTCACACAGATGGTCATTC 20

RESULT 2  
US-09-736-457-981  
; Sequence 981, Application US/09736457  
; Patent No. US20020168637A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong

```

; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darlick
; APPLICANT: Retter, Marc
; APPLICANT: Manion, Jane
; APPLICANT: Fan, Aijun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 981
; LENGTH: 550
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-981

```

```

Query Match      100.0%; Score 20; DB 9; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.26;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 TTCAACCGAGATGTCATTC 20
      |||||
Db      195 TTCAACCGAGATGTCATTC 214

```

```

RESULT 3
US-09-902-941-981
; Sequence 981, Application US/09902941
; Patent No. US20020172952A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tonglong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darlick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.478C17
; CURRENT APPLICATION NUMBER: US/09/902,941
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 2002
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 981
; LENGTH: 550
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-902-941-981

```

```

Query Match      100.0%; Score 20; DB 9; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.26;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 TTCAACCGAGATGTCATTC 20
      |||||
Db      195 TTCAACCGAGATGTCATTC 214

```

```

RESULT 4
US-09-849-626-981
; Sequence 981, Application US/09849626
; Publication No. US20020197669A1

```

```

; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya
; APPLICANT: Fanger, Gary
; APPLICANT: Wang, Aijun
; APPLICANT: Wang, Tonglong
; APPLICANT: Switzer, Anne
; APPLICANT: McNeill, Patricia
; APPLICANT: Clapper, Jonathan
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C16
; CURRENT APPLICATION NUMBER: US/09/849,626
; CURRENT FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 1926
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 981
; LENGTH: 550
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-849-626-981

```

```

Query Match      100.0%; Score 20; DB 9; Length 550;
Best Local Similarity 100.0%; Pred. No. 0.26;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 TTCAACCGAGATGTCATTC 20
      |||||
Db      195 TTCAACCGAGATGTCATTC 214

```

```

RESULT 5
US-09-757-100B-1/C
; Sequence 1, Application US/09757100B
; Patent No. US20010034329A1
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Gaarde, William A.
; APPLICANT: Nero, Pamela S.
; TITLE OF INVENTION: Antisense Modulation of Focal Adhesion Kinase
; FILE REFERENCE: ISPH-0533
; CURRENT APPLICATION NUMBER: US/09/757,100B
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 09/377,310
; PRIOR FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: PCT/US00/18999
; PRIOR FILING DATE: 2000-07-13
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3791
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (233)..(391)
; PUBLICATION INFORMATION:
; JOURNAL: DNA
; VOLUME: 12
; ISSUE: 9
; PAGES: 823-830
; DATE: 1993-11
; DATABASE ACCESSION NUMBER: L13616/Genbank
; DATABASE ENTRY DATE: 1995-01-02
US-09-757-100B-1

```

```

Query Match      100.0%; Score 20; DB 10; Length 3791;
Best Local Similarity 100.0%; Pred. No. 0.39;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 TTCAACCGAGATGTCATTC 20
      |||||
Db      3529 TTCAACCGAGATGTCATTC 3510

```

RESULT 6  
US-09-864-761-21875  
; Sequence 21875, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aeonica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 21875  
; LENGTH: 244  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AL121654.1  
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.97  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.84  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1  
; OTHER INFORMATION: SWISSPROT HIT: P47144, EVALUE 3.50e-02  
US-09-864-761-21875

Query Match 84.0%; Score 16.8; DB 10; Length 244;  
Best Local Similarity 90.0%; Pred. No. 10;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 TTTCACACGATGTCATTC 20  
Db 118 TTTCACACGATGTCATTC 137

RESULT 7  
US-09-864-761-5154  
; Sequence 5154, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aeonica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 5154  
; LENGTH: 400  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AL121654.1  
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.97  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.84  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1  
US-09-864-761-5154

Query Match 84.0%; Score 16.8; DB 10; Length 400;  
Best Local Similarity 90.0%; Pred. No. 11;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 TTTCACACGATGTCATTC 20  
Db 344 TTTCACACGATGTCATTC 363

## RESULT 8

US-09-864-761-16204  
; Sequence 16204, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY  
; FILE REFERENCE: Aecmice-x-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 16204  
; LENGTH: 528  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AL121654.1  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 12  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 15  
US-09-864-761-16204

## Query Match

Best Local Similarity 84.0%; Score 16.8; DB 10; Length 528;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TTTCAACGAGATGTCATTC 20

DB 344 TTTCAACGAGATGTCATTC 363

## RESULT 9

## US-09-866-510-23/c

; Sequence 23, Application US/09866510  
; Patent No. US2002011304A1  
; GENERAL INFORMATION:  
; APPLICANT: KAZLAUSKAS, ANDRIUS  
; APPLICANT: IKUNO, YASUSHI  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES  
; FILE REFERENCE: ERM-104.01  
; CURRENT APPLICATION NUMBER: US/09/866,510  
; CURRENT FILING DATE: 2001-05-25  
; PRIOR APPLICATION NUMBER: 60/250,747  
; PRIOR FILING DATE: 2000-12-01  
; PRIOR APPLICATION NUMBER: 60/289,103  
; PRIOR FILING DATE: 2001-05-07  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 23  
; LENGTH: 1683  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1683)  
US-09-866-510-23

## Query Match

Best Local Similarity 84.0%; Score 16.8; DB 10; Length 1683;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TTTCAACGAGATGTCATTC 20

DB 910 TTTCACTCGATGTCATTC 891

## RESULT 10

US-09-866-510-13/c  
; Sequence 13, Application US/09866510  
; Patent No. US2002011304A1  
; GENERAL INFORMATION:  
; APPLICANT: KAZLAUSKAS, ANDRIUS  
; APPLICANT: IKUNO, YASUSHI  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES  
; FILE REFERENCE: ERM-104.01  
; CURRENT APPLICATION NUMBER: US/09/866,510  
; CURRENT FILING DATE: 2001-05-25  
; PRIOR APPLICATION NUMBER: 60/250,747  
; PRIOR FILING DATE: 2000-12-01  
; PRIOR APPLICATION NUMBER: 60/289,103  
; PRIOR FILING DATE: 2001-05-07  
; NUMBER OF SEQ ID NOS: 33  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 3321  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(3270)  
US-09-866-510-13

## Query Match

Best Local Similarity 84.0%; Score 16.8; DB 10; Length 3321;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TTTCAACGAGATGTCATTC 20

DB 910 TTTCACTCGATGTCATTC 891

## RESULT 11

US-09-866-510-15/c  
; Sequence 15, Application US/09866510  
; Patent No. US2002011304A1

```
; GENERAL INFORMATION:
; APPLICANT: KAZLAUSKAS, ANDRIUS
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES
; FILE REFERENCE: ERM-104.01
; CURRENT FILING DATE: 2001-05-25
; CURRENT APPLICATION NUMBER: US/09/866,510
; PRIOR APPLICATION NUMBER: 60/250,747
; PRIOR FILING DATE: 2000-12-01
; PRIOR FILING DATE: 2000-12-01
; PRIOR FILING DATE: 2000-12-01
; PRIOR FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 3321
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3318)
US-09-866-510-15

Query Match      84.0%; Score 16.8; DB 10; Length 3321;
Best Local Similarity 90.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTCAACACGATGTCATTC 20
||||| ||||| ||||| |||||
DB 910 TTTCATCCTGATGGTCATTC 891

RESULT 12
US-09-866-510-17/c
; Sequence 17, Application US/09866510
; Patent No. US20020111304A1
; GENERAL INFORMATION:
; APPLICANT: KAZLAUSKAS, ANDRIUS
; APPLICANT: IKUNO, YASUSHI
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES
; FILE REFERENCE: ERM-104.01
; CURRENT FILING DATE: 2001-05-25
; CURRENT APPLICATION NUMBER: US/09/866,510
; PRIOR APPLICATION NUMBER: 60/250,747
; PRIOR FILING DATE: 2000-12-01
; PRIOR FILING DATE: 2000-12-01
; PRIOR FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 3321
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3318)
US-09-866-510-17

Query Match      84.0%; Score 16.8; DB 10; Length 3321;
Best Local Similarity 90.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTCAACACGATGTCATTC 20
||||| ||||| ||||| |||||
DB 910 TTTCATCCTGATGGTCATTC 891

RESULT 13
US-09-866-510-19/c
; Sequence 19, Application US/09866510
; Patent No. US20020111304A1
; GENERAL INFORMATION:
; APPLICANT: KAZLAUSKAS, ANDRIUS
; APPLICANT: IKUNO, YASUSHI
```

```
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES
; FILE REFERENCE: ERM-104.01
; CURRENT APPLICATION NUMBER: US/09/866,510
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/250,747
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 60/289,103
; PRIOR FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 3321
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3318)
US-09-866-510-19

Query Match      84.0%; Score 16.8; DB 10; Length 3321;
Best Local Similarity 90.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTCAACACGATGTCATTC 20
||||| ||||| ||||| |||||
DB 910 TTTCATCCTGATGGTCATTC 891

RESULT 14
US-09-866-510-21/c
; Sequence 21, Application US/09866510
; Patent No. US20020111304A1
; GENERAL INFORMATION:
; APPLICANT: KAZLAUSKAS, ANDRIUS
; APPLICANT: IKUNO, YASUSHI
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING FIBROTIC DISEASES
; FILE REFERENCE: ERM-104.01
; CURRENT APPLICATION NUMBER: US/09/866,510
; CURRENT FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/250,747
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: 60/289,103
; PRIOR FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 3321
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(3318)
US-09-866-510-21

Query Match      84.0%; Score 16.8; DB 10; Length 3321;
Best Local Similarity 90.0%; Pred. No. 18;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTCAACACGATGTCATTC 20
||||| ||||| ||||| |||||
DB 910 TTTCATCCTGATGGTCATTC 891

RESULT 15
US-09-955-363-1/c
; Sequence 1, Application US/09955363
; Patent No. US20020173621A1
; GENERAL INFORMATION:
; APPLICANT: Sledziewski Ph.D., Andrzej Z
; Bell, Lillian A.
; Kindsvogel Ph.D., Wayne R.
; TITLE OF INVENTION: METHODS OF PRODUCING SECRETED RECEPTOR ANALOGS
; AND BIOLOGICALLY ACTIVE DIMERIZED POLYPEPTIDE
```

```

1      NUMBER OF SEQUENCES: 36
2      FUSIONS
3      CORRESPONDENCE ADDRESS:
4      ADDRESSEE: Seed and Berry
5      STREET: 6300 Columbia Center
6      CITY: Seattle
7      STATE: WA
8      COUNTRY: USA
9      ZIP: 98104-7092
10     COMPUTER READABLE FORM:
11     MEDIUM TYPE: Floppy disk
12     COMPUTER: IBM PC compatible
13     OPERATING SYSTEM: PC-DOS/MS-DOS
14     SOFTWARE: PatentIn Release #1.24
15     CURRENT APPLICATION DATA:
16     APPLICATION NUMBER: US/09/955,363
17     FILING DATE: 18-Sep-2001
18     CLASSIFICATION: <Unknown>
19     PRIOR APPLICATION DATA:
20     APPLICATION NUMBER: US 07/634,510
21     FILING DATE: <Unknown>
22     APPLICATION NUMBER: US 07/347,291
23     FILING DATE: 02-MAY-1989
24     ATTORNEY/AGENT INFORMATION:
25     NAME: Maki, J.D., David J.
26     REGISTRATION NUMBER: 31,392
27     REFERENCE/DOCKET NUMBER: 990008.446C3
28     TELECOMMUNICATION INFORMATION:
29     TELEPHONE: 206-622-4900
30     TELEFAX: 206-682-6031
31     TELEX: 3723836
32     INFORMATION FOR SEQ ID NO: 1:
33     .SEQUENCE CHARACTERISTICS:
34     LENGTH: 4465 base pairs
35     TYPE: nucleic acid
36     STRANDEDNESS: double
37     TOPOLOGY: linear
38     MOLECULE TYPE: cDNA
39     HYPOTHETICAL: N
40     ANTI-SENSE: N
41     ORIGINAL SOURCE:
42     ORGANISM: Homo sapiens
43     DEVELOPMENTAL STAGE: Adult
44     TISSUE TYPE: Skin
45     CELL TYPE: fibroblasts
46     IMMEDIATE SOURCE:
47     CLONE: pr-rx1
48     FEATURE:
49     NAME/KEY: CDS
50     LOCATION: 354..3671
51     OTHER INFORMATION:
52     SEQUENCE DESCRIPTION: SEQ ID NO: 1:
53     ;
54     ;
55     ;
56     ;
57     ;
58     ;
59     ;
60     ;
61     ;
62     ;
63     ;
64     ;
65     ;
66     ;
67     ;
68     ;
69     ;
70     ;
71     ;
72     ;
73     ;
74     ;
75     ;
76     ;
77     ;
78     ;
79     ;
80     ;
81     ;
82     ;
83     ;
84     ;
85     ;
86     ;
87     ;
88     ;
89     ;
90     ;
91     ;
92     ;
93     ;
94     ;
95     ;
96     ;
97     ;
98     ;
99     ;
100    ;
101    ;
102    ;
103    ;
104    ;
105    ;
106    ;
107    ;
108    ;
109    ;
110    ;
111    ;
112    ;
113    ;
114    ;
115    ;
116    ;
117    ;
118    ;
119    ;
120    ;
121    ;
122    ;
123    ;
124    ;
125    ;
126    ;
127    ;
128    ;
129    ;
130    ;
131    ;
132    ;
133    ;
134    ;
135    ;
136    ;
137    ;
138    ;
139    ;
140    ;
141    ;
142    ;
143    ;
144    ;
145    ;
146    ;
147    ;
148    ;
149    ;
150    ;
151    ;
152    ;
153    ;
154    ;
155    ;
156    ;
157    ;
158    ;
159    ;
160    ;
161    ;
162    ;
163    ;
164    ;
165    ;
166    ;
167    ;
168    ;
169    ;
170    ;
171    ;
172    ;
173    ;
174    ;
175    ;
176    ;
177    ;
178    ;
179    ;
180    ;
181    ;
182    ;
183    ;
184    ;
185    ;
186    ;
187    ;
188    ;
189    ;
190    ;
191    ;
192    ;
193    ;
194    ;
195    ;
196    ;
197    ;
198    ;
199    ;
200    ;
201    ;
202    ;
203    ;
204    ;
205    ;
206    ;
207    ;
208    ;
209    ;
210    ;
211    ;
212    ;
213    ;
214    ;
215    ;
216    ;
217    ;
218    ;
219    ;
220    ;
221    ;
222    ;
223    ;
224    ;
225    ;
226    ;
227    ;
228    ;
229    ;
230    ;
231    ;
232    ;
233    ;
234    ;
235    ;
236    ;
237    ;
238    ;
239    ;
240    ;
241    ;
242    ;
243    ;
244    ;
245    ;
246    ;
247    ;
248    ;
249    ;
250    ;
251    ;
252    ;
253    ;
254    ;
255    ;
256    ;
257    ;
258    ;
259    ;
260    ;
261    ;
262    ;
263    ;
264    ;
265    ;
266    ;
267    ;
268    ;
269    ;
270    ;
271    ;
272    ;
273    ;
274    ;
275    ;
276    ;
277    ;
278    ;
279    ;
280    ;
281    ;
282    ;
283    ;
284    ;
285    ;
286    ;
287    ;
288    ;
289    ;
290    ;
291    ;
292    ;
293    ;
294    ;
295    ;
296    ;
297    ;
298    ;
299    ;
300    ;
301    ;
302    ;
303    ;
304    ;
305    ;
306    ;
307    ;
308    ;
309    ;
310    ;
311    ;
312    ;
313    ;
314    ;
315    ;
316    ;
317    ;
318    ;
319    ;
320    ;
321    ;
322    ;
323    ;
324    ;
325    ;
326    ;
327    ;
328    ;
329    ;
330    ;
331    ;
332    ;
333    ;
334    ;
335    ;
336    ;
337    ;
338    ;
339    ;
340    ;
341    ;
342    ;
343    ;
344    ;
345    ;
346    ;
347    ;
348    ;
349    ;
350    ;
351    ;
352    ;
353    ;
354    ;
355    ;
356    ;
357    ;
358    ;
359    ;
360    ;
361    ;
362    ;
363    ;
364    ;
365    ;
366    ;
367    ;
368    ;
369    ;
370    ;
371    ;
372    ;
373    ;
374    ;
375    ;
376    ;
377    ;
378    ;
379    ;
380    ;
381    ;
382    ;
383    ;
384    ;
385    ;
386    ;
387    ;
388    ;
389    ;
390    ;
391    ;
392    ;
393    ;
394    ;
395    ;
396    ;
397    ;
398    ;
399    ;
400    ;
401    ;
402    ;
403    ;
404    ;
405    ;
406    ;
407    ;
408    ;
409    ;
410    ;
411    ;
412    ;
413    ;
414    ;
415    ;
416    ;
417    ;
418    ;
419    ;
420    ;
421    ;
422    ;
423    ;
424    ;
425    ;
426    ;
427    ;
428    ;
429    ;
430    ;
431    ;
432    ;
433    ;
434    ;
435    ;
436    ;
437    ;
438    ;
439    ;
440    ;
441    ;
442    ;
443    ;
444    ;
445    ;
446    ;
447    ;
448    ;
449    ;
450    ;
451    ;
452    ;
453    ;
454    ;
455    ;
456    ;
457    ;
458    ;
459    ;
460    ;
461    ;
462    ;
463    ;
464    ;
465    ;
466    ;
467    ;
468    ;
469    ;
470    ;
471    ;
472    ;
473    ;
474    ;
475    ;
476    ;
477    ;
478    ;
479    ;
480    ;
481    ;
482    ;
483    ;
484    ;
485    ;
486    ;
487    ;
488    ;
489    ;
490    ;
491    ;
492    ;
493    ;
494    ;
495    ;
496    ;
497    ;
498    ;
499    ;
500    ;
501    ;
502    ;
503    ;
504    ;
505    ;
506    ;
507    ;
508    ;
509    ;
510    ;
511    ;
512    ;
513    ;
514    ;
515    ;
516    ;
517    ;
518    ;
519    ;
520    ;
521    ;
522    ;
523    ;
524    ;
525    ;
526    ;
527    ;
528    ;
529    ;
530    ;
531    ;
532    ;
533    ;
534    ;
535    ;
536    ;
537    ;
538    ;
539    ;
540    ;
541    ;
542    ;
543    ;
544    ;
545    ;
546    ;
547    ;
548    ;
549    ;
550    ;
551    ;
552    ;
553    ;
554    ;
555    ;
556    ;
557    ;
558    ;
559    ;
560    ;
561    ;
562    ;
563    ;
564    ;
565    ;
566    ;
567    ;
568    ;
569    ;
570    ;
571    ;
572    ;
573    ;
574    ;
575    ;
576    ;
577    ;
578    ;
579    ;
580    ;
581    ;
582    ;
583    ;
584    ;
585    ;
586    ;
587    ;
588    ;
589    ;
590    ;
591    ;
592    ;
593    ;
594    ;
595    ;
596    ;
597    ;
598    ;
599    ;
600    ;
601    ;
602    ;
603    ;
604    ;
605    ;
606    ;
607    ;
608    ;
609    ;
610    ;
611    ;
612    ;
613    ;
614    ;
615    ;
616    ;
617    ;
618    ;
619    ;
620    ;
621    ;
622    ;
623    ;
624    ;
625    ;
626    ;
627    ;
628    ;
629    ;
630    ;
631    ;
632    ;
633    ;
634    ;
635    ;
636    ;
637    ;
638   
```

Search completed: January 12, 2003, 12:08:34  
Job time : 59 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 12, 2003, 09:45:15 ; Search time 1993 Seconds  
(without alignments)  
162.524 Million cell updates/sec

Title: US-09-757-100B-18

Perfect score: 20  
Sequence: 1 tttaaccagtggtcattc 20

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*

- 1: em\_estba:\*
- 2: em\_esthum:\*
- 3: em\_estin:\*
- 4: em\_estnu:\*
- 5: em\_estov:\*
- 6: em\_estpl:\*
- 7: em\_estro:\*
- 8: em\_htc:\*
- 9: gb\_est1:\*
- 10: gb\_est2:\*
- 11: gb\_htc:\*
- 12: gb\_est3:\*
- 13: gb\_est4:\*
- 14: gb\_est5:\*
- 15: em\_estfun:\*
- 16: em\_estom:\*
- 17: gb\_gss:\*
- 18: em\_gss\_hum:\*
- 19: em\_gss\_inv:\*
- 20: em\_gss\_pin:\*
- 21: em\_gss\_vrt:\*
- 22: em\_gss\_fun:\*
- 23: em\_gss\_mam:\*
- 24: em\_gss\_mus:\*
- 25: em\_gss\_other:\*
- 26: em\_gss\_pro:\*
- 27: em\_gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length	Description
C 1	20	100.0	285 12 BF997449
C 2	20	100.0	351 14 R32551
C 3	20	100.0	392 10 AV709841
C 4	20	100.0	412 10 BE092864
C 5	20	100.0	416 14 R89170
C 6	20	100.0	424 10 AV724065

C	7	20	100.0	425	10	BE566855
C	8	20	100.0	429	10	AW136503
C	9	20	100.0	432	9	AA983514
C	10	20	100.0	447	13	BM664187
C	11	20	100.0	447	14	N42230
C	12	20	100.0	461	9	AI341506
C	13	20	100.0	481	12	BF084847
C	14	20	100.0	492	13	BG986106
C	15	20	100.0	527	9	AI016118
C	16	20	100.0	532	9	AA404694
C	17	20	100.0	569	9	AI380178
C	18	20	100.0	585	14	BM790849
C	19	20	100.0	598	9	AA453479
C	20	20	100.0	610	10	AW003490
C	21	20	100.0	632	12	BF669338
C	22	20	100.0	648	10	AW992843
C	23	20	100.0	648	17	AG060293
C	24	20	100.0	676	12	BF691197
C	25	20	100.0	682	10	AW992847
C	26	20	100.0	690	17	AQ424326
C	27	20	100.0	693	12	BF978799
C	28	20	100.0	716	12	BF701452
C	29	20	100.0	722	10	BE567556
C	30	20	100.0	769	14	BQ438629
C	31	20	100.0	771	10	AV726076
C	32	20	100.0	808	10	BE380127
C	33	20	100.0	824	12	BF665639
C	34	20	100.0	849	12	BF241766
C	35	20	100.0	849	12	BG528906
C	36	20	100.0	852	10	BE617048
C	37	20	100.0	872	12	BF675982
C	38	20	100.0	882	12	BF217100
C	39	20	100.0	885	12	BF211871
C	40	20	100.0	888	12	BF698361
C	41	20	100.0	896	12	BF690762
C	42	20	100.0	906	12	BF213328
C	43	20	100.0	908	12	EG110386
C	44	20	100.0	930	12	BG617438
C	45	20	100.0	972	13	BI917958

ALIGNMENTS

RESULT 1  
BF997449/c

LOCUS BF997449 285 bp mRNA linear EST 23-JAN-2001  
DEFINITION CM0-GN0100-081100-576-a05 GN0100 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BF997449  
VERSION BF997449.1 GI:12403772  
KEYWORDS EST.  
SOURCE human.

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 285)  
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,  
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare  
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
MEDLINE 20202663  
COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001

Email: asimpson@ludwig.org.br  
 This sequence was derived from the PAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?cl=CM0&c2=CM0-GN0100-081100-576-a05&c3=2000-11-08&c4=1)  
 Seq primer: puc18 forward  
 High quality sequence start: 10  
 High quality sequence stop: 285.  
 Location/Qualifiers

## FEATURES

source

1..285

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone\_lib="GN0100"  
 /dev\_stage="Adult"

/note="Organ: placenta\_normal; Vector: puc18; Site\_1: SmaI; Site\_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

## BASE COUNT

71 a 69 c 46 g 99 t

## ORIGIN

Query Match 100.0%; Score 20; DB 12; Length 285;  
 Best Local Similarity 100.0%; Pred. No. 21;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## QY

1 TTTCACCGAGATGTCATTC 20

Db 163 TTTCACCGAGATGTCATTC 144

RESULT 2  
 R32551/c 351 bp mRNA linear EST 28-APR-1995  
 LOCUS yH5&d07.r1 Soares placenta NB2HP homo sapiens cDNA clone  
 DEFINITION IMAGE:133549.5' similar to gb:L05186 FOCAL ADHESION KINASE (HUMAN); mRNA sequence.

ACCESSION R32551  
 VERSION R32551.1 GI:788394  
 KEYWORDS EST  
 SOURCE human.

## ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 351)  
 Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M., Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M., Parsons J.,

Ritkin L., Rohlfing T., Soares M., Tan F., Trevaskis E., Waterston R., Williamson A., Wohldmann P. and Wilson R.  
 The WashU-Merck EST Project  
 Unpublished (1995)

## TITLE

Contact: Wilson R.

Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu

## JOURNAL

High quality sequence stop: 299

Source: IMAGE Consortium, LNL  
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.  
 Insert length: 691 Std Error: 0.00  
 Seq primer: M13R1  
 High quality sequence stop: 299.

## FEATURES

source

1..351

/organism="Homo sapiens"  
 /db\_xref="GDB:539275"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:133549"

/clone\_lib="Soares placenta NB2HP"  
 /sex="Female"  
 /dev\_stage="placenta obtained at birth (full term)"  
 /lab\_host="DH10B (ampicillin resistant)"  
 /note="Organ: placenta; Vector: pT73D (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' AACTGAGAGATTGCGCGCCGACGAGATTGTTTCTTTTCTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."

## BASE COUNT

100 a 66 c 65 g 120 t

## ORIGIN

Query Match 100.0%; Score 20; DB 14; Length 351;  
 Best Local Similarity 100.0%; Pred. No. 24;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## QY

1 TTTCACCGAGATGTCATTC 20

Db 104 TTTCACCGAGATGTCATTC 85

RESULT 3  
 AV709841/c 392 bp mRNA linear EST 09-OCT-2000  
 LOCUS AV709841 ADC Homo sapiens cDNA clone ADCAG11 5', mRNA sequence.

ACCESSION AV709841  
 VERSION AV709841.1 GI:10727915  
 KEYWORDS EST.  
 SOURCE human.

## ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 392)  
 Peng Y., Song H., Huang Q., Huang C., Gu Y., Yang Y., Gao G., Xiao H., Xu X., Li N., Qian B., Liu F., Qu J., Gao X., Cheng Z., Xu Z.,

Zeng L., Xu S., Gu M., Tu Y., Jia J., Fu G., Ren S., Zhong M., Lu G., Hu R., Chen J., Chen Z. and Han Z.  
 Homo sapiens cDNA ADC clones  
 Unpublished (2000)

## JOURNAL

Contact: Zeguang Han

Chinese National Human Genome Center at Shanghai  
 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
 201203, P. R. China  
 Tel: 86-21-50801919 (ex. 45)  
 Fax: 86-21-50801922

Email: hanzg@chgc.sh.cn  
 This clone is available at CHGC in Shanghai.  
 Location/Qualifiers

## FEATURES

1..392

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="ADCAG11"

/clone\_lib="ADC"  
 /tissue\_type="Adrenal gland"  
 /dev\_stage="Adult"  
 /lab\_host="SOLR"

/note="Vector: pBluescript sk(-); Site\_1: EcoRI; Site\_2: XhoI"

BASE COUNT 107 a 85 c 82 g 116 t 2 others

## ORIGIN

Query Match 100.0%; Score 20; DB 10; Length 392;  
 Best Local Similarity 100.0%; Pred. No. 25;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## QY

1 TTTCACCGAGATGTCATTC 20

Db 173 TTTCACCGAGATGTCATTC 154



COMMENT Contact: Zeguang Han  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
201203, P. R. China  
Tel: 86-21-50801919(ex.45)  
Fax: 86-21-50801922  
Email: hanzg@chgc.sh.cn  
This clone is available at CHGC in Shanghai.

FEATURES  
source  
1..424  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="HTBA1F11"  
/clone\_1fb="HTB"  
/tissue\_type="Hypothalamus"  
/dev\_stage="Adult"  
/lab\_host="SOLR"  
/note="Vector: pBluescript sk(-); Site\_1: EcoRI; Site\_2: XhoI"

BASE COUNT 118 a 90 c 80 g 136 t

ORIGIN

Query Match 100.0%; Score 20; DB 10; Length 424;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTTCAACCGAGTGGTCATTC 20  
|||||  
DB 215 TTTCAACCGAGTGGTCATTC 196

RESULT 7  
BE566855/c 425 bp mRNA linear EST 15-AUG-2000  
LOCUS 601339461F1 NIH\_MGC\_53 Homo sapiens cDNA clone IMAGE:3681878 5',  
DEFINITION mRNA sequence.  
ACCESSION BE566855  
VERSION BE566855.1 GI:9810575  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 425)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue procurement: ATCC  
CDNA library preparation: CLONTECH Laboratories, Inc.  
DNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at: [image.llnl.gov](http://image.llnl.gov)  
Plate: LLC364 row: 1 column: 15  
High quality sequence scop: 421.  
Location/Qualifiers  
1..425  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3681878"  
/clone\_1fb="NIH\_MGC\_53"  
/tissue\_type="Carcinoma, cell line"  
/lab\_host="DH10B (T1 phage-resistant)"  
/note="Organ: bladder; Vector: pDNR-LIB (Clontech);  
Site\_1: SfiI (ggcgccgagcgc); Site\_2: SfiI (ggccatctggcc  
); Double-stranded cDNA was prepared from cell line RNA.  
5' and 3' adaptors were used in cloning as follows: 5'  
adaptor sequence: 5'-CACGCCATATGCGC-3' and 3' adaptor  
sequence: 5'-ATTCTAGAGCGGAGCGCGCGCATG-3' (30 BN-3'  
(where B = A, C, or G and N = A, C, G, or T) Average  
insert size 1.55 kb (range 0.9-4.0 kb). 15/15 colonies

contained inserts by PCR. This library was enriched for  
full-length clones and was constructed by Clontech  
Laboratories (Palo Alto, CA)."

BASE COUNT 117 a 118 c 91 g 99 t

ORIGIN

Query Match 100.0%; Score 20; DB 10; Length 425;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTTCAACCGAGTGGTCATTC 20  
|||||  
DB 376 TTTCAACCGAGTGGTCATTC 357

RESULT 8  
AW136503 429 bp mRNA linear EST 29-OCT-1999  
LOCUS UI-H-B11-acx-e-10-0-UI.s1 NCI\_CGAP\_Sub3 Homo sapiens cDNA clone  
DEFINITION IMAGE:2715763 3', mRNA sequence.  
ACCESSION AW136503  
VERSION AW136503.1 GI:6140563  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 429)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
The sequence contained an oligo-dT track that was present in the  
oligonucleotide that was used to prime the synthesis of first  
strand cDNA and therefore this may represent a bonafide poly A  
tail. CDNA library preparation: M.B. Soares Lab Clone distribution:  
NCI-CGAP clone distribution information can be found through the  
I.M.A.G.E. Consortium/LLNL at:  
[www-bio.llnl.gov/bdtp/image/image.html](http://www-bio.llnl.gov/bdtp/image/image.html)  
Seq primer: M13 Forward  
POLYA=Yes.  
Location/Qualifiers  
1..429  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2715763"  
/clone\_1fb="NCI\_CGAP\_Sub3"  
/lab\_host="DH10B (Life Technologies)"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; Site\_1: Not I; Site\_2: Eco RI; The  
NCI\_CGAP Sub3 library is a subtracted library derived from  
the NCI\_CGAP Sub1 library, which is a subtracted library  
derived from B1. B1 constitutes a mixture of 21  
normalized or subtracted NCI\_CGAP libraries: NCI\_CGAP\_C04  
, NCI\_CGAP\_P122, NCI\_CGAP\_P128, NCI\_CGAP\_C010,  
NCI\_CGAP\_C016, NCI\_CGAP\_K105, NCI\_CGAP\_K1d12,  
NCI\_CGAP\_K1d3, NCI\_CGAP\_K1d11, NCI\_CGAP\_Lym2,  
NCI\_CGAP\_Br2, NCI\_CGAP\_C08, NCI\_CGAP\_C1u1, NCI\_CGAP\_L2i2,  
NCI\_CGAP\_Br23, NCI\_CGAP\_Lu5, NCI\_CGAP\_Lu24,  
NCI\_CGAP\_Lu19, NCI\_CGAP\_GC4, NCI\_CGAP\_GC6,  
NCI\_CGAP\_Brn25. These 21 libraries were pooled and a  
single-stranded DNA preparation of the resulting mixture  
was used as a tracer in a subtractive hybridization with  
a driver whose composition is detailed below:  
NCI\_CGAP\_K1d3 pool 1 LLM 3334-3337, 3682-3683,  
3798-3803 (IMAGE Clonides 1322376-1323911, 1456008-1456775  
, 1500552-1502855); NCI\_CGAP\_K1d5 pool 1 LLM 3338-3342  
, 3722-3725, 3776-3778 (IMAGE Clonides 1323912-1325831,  
1471368-1472903, 1492104-1493255); NCI\_CGAP\_Lu5 pool 1  
LLM 3575-3582, 3851-3854 (IMAGE Clonides 1414920-1417991,  
1520904-1524239); NCI\_CGAP\_GC4 pool 1 LLM 3164-3167,

3716-3720, 3733-3735 (IMAGE Clones) 1257096-1258631,  
1459054-1470983, 1475552-1476743; NCI\_CGAP\_P22 pool 1  
LLAM 2457-2459, 2758-2759, 3062-3068 (IMAGE Clones)  
985608-986759, 1101192-1101959, 1217928-1220615);  
NCI CGAP Co10 pool 1 LLAM 2644-2653, 2871-2872 (IMAGE  
Clones) 1057416-1061255, 1144584-1145351). Subtraction  
was performed as previously described [Bonaldo, Lennon &  
Soares (1996): Normalization and Subtraction: Two  
Approaches to Facilitate Gene Discovery. Genome Research  
6, 791-806.  
TAG LIB=NCI CGAP GC4  
TAG TISSUE=Germ Cell  
TAG\_SEQ=AAATC" 113 a 80 c 105 g 131 t

BASE COUNT  
ORIGIN

Query Match 100.0%; Score 20; DB 10; Length 429;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TTTCAACACGATGTCATTC 20  
|||||  
Db 108 TTTCAACACGATGTCATTC 127  
|||||

RESULT 9  
AA983514 432 bp mRNA linear EST 27-AUG-1998  
LOCUS or41g10.s1 NCI CGAP GC3 Homo sapiens cDNA clone IMAGE:1598466 3'  
DEFINITION similar to gb:L05186 FOCAL ADHESION KINASE (HUMAN);, mRNA sequence.  
ACCESSION AA983514  
VERSION  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 432)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaps-femail.nih.gov  
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael  
Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D.  
DNA Sequencing by: Greg Lennon, Ph.D.  
Cloning Distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www.bio.llnl.gov/bbrp/image/image.html  
Insert Length: 1419 Std Error: 0.00  
Seq primer: -40m13 fwd. RT from Amersham.

FEATURES  
source

1. 432  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1598466"  
/clone\_lib="NCI\_CGAP\_GC3"  
/tissue\_type="pooled germ cell tumors"  
/lab\_host="DH108"  
/notes="Vector: pT7T3D-Pac (Pharmacia) with a modified  
polylinker; 1st strand cDNA was prepared from 3 pooled  
germ cell tumors, and was then primed with a Not I -  
oligo(dT) primer. Double-stranded cDNA was ligated to Eco  
RI adaptors (Pharmacia), digested with Not I and cloned  
into the Not I and Eco RI sites of the modified pT7T3  
vector. Library is not normalized. Library was  
constructed by Bento Soares and M. Fatima Bonaldo."  
116 a 81 c 114 g 121 t

BASE COUNT  
ORIGIN

Query Match 100.0%; Score 20; DB 9; Length 432;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TTTCAACACGATGTCATTC 20  
|||||  
Db 91 TTTCAACACGATGTCATTC 110  
|||||

RESULT 10  
BM664187 447 bp mRNA linear EST 27-FEB-2002  
LOCUS UI-E-C11-agg-d-10-0-UI.s1 UI-E-C11 Homo sapiens cDNA clone  
DEFINITION UI-E-C11-agg-d-10-0-UI 3', mRNA sequence.  
ACCESSION BM664187  
VERSION  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 447)  
AUTHORS Bonaldo, M.F., Lennon, G. and Soares, M.B.  
TITLE Normalization and subtraction: two approaches to facilitate gene  
discovery  
JOURNAL Genome Res. 6 (9), 791-806 (1996)  
MEDLINE 97044477  
COMMENT Contact: Soares, MB  
Program for Rat Gene Discovery and Mapping  
University of Iowa  
451 Eckstein Medical Research Building Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: msoares@blue.weeg.uiowa.edu  
Tissue Procurement: Dr. Gregg Hageman  
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Cloning Distribution: Researchers may obtain clones from Research  
Genetics (www.resgen.com).  
The following repetitive elements were found in this cDNA  
sequence: 1-22, >AT-rich#Low\_complexity (matched complement)  
Seq primer: M13 Forward  
POLYA=Yes.

FEATURES  
source

1. 447  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="UI-E-C11-agg-d-10-0-UI"  
/clone\_lib="UI-E-C11"  
/tissue\_type="RPE and Choroid"  
/dev\_stage="adult"  
/lab\_host="DH108 (Life Technologies) (T1 phage resistant)"  
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a  
modified polylinker; Site 1: EcoR I; Site 2: Not I;  
UI-E-C11 is a normalized cDNA library containing the  
following tissue(s): RPE and Choroid. The library was  
constructed according to Bonaldo, Lennon and Soares,  
Genome Research, 6:791-806, 1996. First strand cDNA  
synthesis was primed with an oligo-dT primer containing a  
Not I site. Double stranded cDNA was ligated to an EcoR I  
adaptor, digested with Not I, and cloned directionally  
into pT7T3-Pac vector. The oligonucleotide used to prime  
the synthesis of first-strand cDNA contains a library tag  
sequence that is located between the Not I site and the  
(dT)18 tail. The sequence tag for this library is ACCTTA.  
This library was created for the program, Gene Discovery  
in the Visual System, supported by National Eye Institute  
(NEI).  
TAG LIB=UI-E-C11  
TAG\_TISSUE=RPE and Choroid  
TAG\_SEQ=ACCTTA" 127 a 81 c 103 g 136 t

BASE COUNT

## ORIGIN

Query Match 100.0%; Score 20; DB 13; Length 447;  
 Best Local Similarity 100.0%; Pred. No. 27;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTTCAACGAGATGTCATTC 20  
 |||  
 Db 157 TTTCAACGAGATGTCATTC 176

## RESULT 11

LOCUS M42230/c 462 bp mRNA linear EST 24-JAN-1996  
 DEFINITION yw71h06.r1 Soares placenta\_8to9weeks\_2NBHP8to9w Homo sapiens cDNA  
 clone IMAGE:257723 5' similar to gb:I05186 FOCAL ADHESION KINASE  
 (HUMAN);, mRNA sequence.

ACCESSION M42230  
 VERSION M42230.1 GI:1166261  
 KEYWORDS EST.  
 SOURCE human.

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 462)

REFERENCE Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman  
 M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Maiz, M., Parsons, J.,  
 Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevisakis, E., Waterston  
 R., Williamson, A., Wohldmann, P. and Wilson, R.  
 The Washu-Merck EST Project

JOURNAL Unpublished (1995)  
 COMMENT Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810

EMAIL: est@watson.wustl.edu  
 High quality sequence stops: 330  
 Source: IMAGE Consortium, LINT  
 This clone is available royalty-free through LINT; contact the  
 IMAGE Consortium (info@image.lint.gov) for further information.  
 Seq primer: T7  
 High quality sequence stop: 330.  
 Location/Qualifiers

FEATURES  
 source 1..462

/organism="Homo sapiens"  
 /db\_xref="GDB:388733"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:257723"  
 /clone\_lib="Soares\_placenta\_8to9weeks\_2NBHP8to9w"  
 /dev\_stage="two placentae: one from 8 weeks and another  
 from 9 weeks post conception"  
 /lab\_host="DH10B (ampicillin resistant)"  
 /note="organ: placenta; Vector: pRTT3D (Pharmacia) with a  
 modified polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st  
 strand cDNA was primed with a Not I - oligo(dT) primer [5'  
 TGTACCAATCGAAGTGGAGCGCGCGATTTTCTTTTCTTTT 3']  
 TGTACCAATCGAAGTGGAGCGCGCGATTTTCTTTTCTTTT 3']  
 double-stranded cDNA was size selected, ligated to Eco RI  
 adapters (Pharmacia), digested with Not I and cloned into  
 the Not I and Eco RI sites of a modified pRTT3 vector  
 (Pharmacia). Library constructed by Bento Soares and  
 M. Fatima Bonaldo."

BASE COUNT 137 a 96 c 89 g 139 t 1 others  
 ORIGIN

Query Match 100.0%; Score 20; DB 14; Length 462;  
 Best Local Similarity 100.0%; Pred. No. 28;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTTCAACGAGATGTCATTC 20  
 |||  
 Db 193 TTTCAACGAGATGTCATTC 174

## RESULT 12

LOCUS A1341506 481 bp mRNA linear EST 15-FEB-1999  
 DEFINITION qx92a12.x1 NCI CGAP GC6 Homo sapiens cDNA clone IMAGE:200950 3'  
 similar to gb:I05186 FOCAL ADHESION KINASE (HUMAN);, mRNA sequence.

ACCESSION A1341506  
 VERSION A1341506.1 GI:4078433  
 KEYWORDS EST.  
 SOURCE human.

## ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 481)  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index

JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cgaps-remail.nih.gov  
 Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael  
 R. Emmert-Buck, M.D., Ph.D.  
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
 Bonaldo, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LINT at:  
[www-bio.lint.gov/bhrp/image/image.html](http://www-bio.lint.gov/bhrp/image/image.html)  
 Insert length: 686 Std Error: 0.00  
 Seq primer: -40UP from Gibco

High quality sequence stop: 445.  
 Location/Qualifiers

FEATURES  
 source 1..481

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:200950"  
 /clone\_lib="NCI-CGAP\_GC6"  
 /tissue\_type="pooled germ cell tumors"  
 /lab\_host="DH10B"  
 /note="Vector: pRTT3D-Pac (Pharmacia) with a modified  
 polylinker; Site\_1: Not I; Site\_2: Eco RI; Plasmid DNA  
 from the normalized library NCI-CGAP\_GC4 was prepared, and  
 89 circles were made in vitro. Following HAP purification,  
 this DNA was used as tracer in a subtractive hybridization  
 reaction. The driver was PCR-amplified cDNAs from a pool  
 of 5,000 clones made from the same library (cloneids  
 1257096-1258631, 1469064-1470983, and 1475592-1476743).  
 Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 128 a 98 c 123 g 132 t  
 ORIGIN

Query Match 100.0%; Score 20; DB 9; Length 481;  
 Best Local Similarity 100.0%; Pred. No. 28;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTTCAACGAGATGTCATTC 20  
 |||  
 Db 92 TTTCAACGAGATGTCATTC 111

## RESULT 13

LOCUS BF084847/c 492 bp mRNA linear EST 18-OCT-2000  
 DEFINITION PM2-DT0042-100900-003-g03 DT0042 Homo sapiens cDNA, mRNA sequence.

ACCESSION BF084847  
 VERSION BF084847.1 GI:10878677  
 KEYWORDS EST.  
 SOURCE human.

## ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 492)

**AUTHORS**  
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

**TITLE**  
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

**JOURNAL MEDLINE COMMENT**  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663

**CONTACT:** Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=&t2=PM2-DT0042-100>  
900-003-g03&t3=2000-09-10&t4=1)

Seq primer: puc 18 forward  
High quality sequence start: 5  
High quality sequence stop: 492.

**FEATURES**  
source  
1. .492  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="DT0042"  
/dev\_stage="Adult"  
/note="Organ: denis drash; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."  
139 a 114 c 99 g 140 t  
BASE COUNT  
ORIGIN

Query Match 100.0%; Score 20; DB 12; Length 492;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTTCACACAGATGGTCATTC 20  
|||||  
Db 325 TTTCACACAGATGGTCATTC 306

**RESULT 14**  
BG986106/c  
LOCUS BG986106 492 bp mRNA linear EST 13-JUN-2001  
DEFINITION PM2-DT0042-190201-005-g03 DT0042 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BG986106  
VERSION BG986106.1 GI:14390176  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 492)  
REFERENCE Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
20202663  
CONTACT: Simpson A.J.G.

**AUTHORS**  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=PM2&t2=PM2-DT0042-190201-005-g03&t3=2001-02-19&t4=1>)

Seq primer: puc 18 forward  
High quality sequence start: 5  
High quality sequence stop: 492.

**FEATURES**  
source  
1. .492  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="DT0042"  
/dev\_stage="Adult"  
/note="Organ: denis drash; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."  
139 a 114 c 99 g 140 t  
BASE COUNT  
ORIGIN

Query Match 100.0%; Score 20; DB 13; Length 492;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTTCACACAGATGGTCATTC 20  
|||||  
Db 325 TTTCACACAGATGGTCATTC 306

**RESULT 15**  
AI016118  
LOCUS AI016118 527 bp mRNA linear EST 27-AUG-1998  
DEFINITION ct34g07.s1 Soares testis NHT Homo sapiens cDNA clone IMAGE:1618716 3' similar to gb:L05186 FOCAL ADHESION KINASE (HUMAN);, mRNA sequence.  
ACCESSION AI016118  
VERSION AI016118.1 GI:3230454  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 527)  
REFERENCE NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: [www-bio.llnl.gov/bbrp/image/image.html](http://www-bio.llnl.gov/bbrp/image/image.html)  
Insert Length: 752 Std Error: 0.00  
Seq primer: -40m13 fwd. ET from Amersham  
High quality sequence stop: 344.  
Location/Qualifiers  
1. .527  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"

```

/clone="IMAGE:1618716"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="Vector: pT773D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I - oligo(dT) primer [5'
TGTACCAATCTGAGATGGAGCGCGCCCAATTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo."

```

```

BASE COUNT      135 a      104 c      134 g      154 t
ORIGIN

```

```

Query Match      100.0%; Score 20; DB 9; Length 527;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 TTTCACCAAGATGTCATTC 20
        |||||
Db      100 TTTCACCAAGATGTCATTC 119

```

Search completed: January 12, 2003, 11:09:54  
Job time : 1997 secs